Mathematical Modeling and AI
5th year

Poisson processes and application to reliability theory and actuarial science

Academic year 2023-2024

Mélisande ALBERT

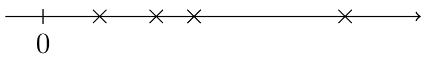
Office 115 GMM
melisande.albert@insa-toulouse.fr

Contents

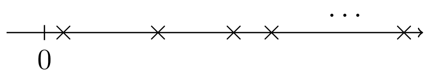
Introduction

The Poisson process, named after the French mathematician Siméon Denis Poisson, lies among the most important stochastic processes in probability theory. It is widely used to model random "points" in time and space, such as radioactive emission times, the arrival times of customers at a service center, earthquakes, eye movements or the positions of flaws in a piece of material.

The Poisson process is used as a foundation for building a number of other, more complicated random processes (e.g. the compound Poisson process). The analysis of such processes differs from usual real or vectorial frameworks and has a particular mathematical formalism.



Usual vectorial framework $(\mathbb{R}_+)^n$ case: fixed number of observations



Point process framework: random number of observations

The aim of this lecture is to introduce such processes in time and to study their basic properties form both probabilistic and statistical aspects. Application to reliability theory and to actuarial sciences will be done in some examples.

- **Application to reliability theory.** We consider a system subject to recurrent failures. We assume that when the system fails, it is repaired to a functioning state, and that the repair time is negligible. Different questions arise:
  - How can we model the successive failure times? How are they distributed in time?
  - Are the failure occurrences stable?
  - Can we estimate the rate of occurence failure per unit of time?
  - ...
- **Application to actuarial sciences.** We aim at describing and modeling insurer's vulnerability to ruin. Claims happen at random times, with random sizes.
  - How are distributed the claim times? and the amounts?
  - How can we model the total claim amount or the risk of failure?
  - Can we estimate the ruin probability?
  - How should the insurance company calibrate the premium?
  - ...

Abbreviations

a.s.	almost surely
c.d.f	cumulative distribution function
CLT	Central Limit Theorem
i.i.d.	independent and identically distributed
LLN	Law of Large Numbers
m.g.f.	moment-generating function
MLE	Maximum Likelihood Estimator
r.v.	random variable
s.t.	such that
w.r.t	with respect to

Chapter 1

Some reminders in probability

Several important probability distributions arise naturally from the Poisson process: the Poisson distribution, the exponential distribution, and the gamma distribution. This chapter aims to recall basic definitions and properties in probability theory necessary to define and study Poisson processes.

I Poisson random variable

The Poisson distribution is a discrete probability distribution commonly used to model the number of (rare) events occurring in a fixed interval of time or space. For instance, an insurer might want to model the number of claims during a month, or in quality control, one might want to model the number of failures of a manufacturing machine during a week.

Definition 1.1.
Let N be a real random variable. Then N is said to have a Poisson distribution with parameter $\lambda > 0$, denoted $N \sim \mathcal{P}(\lambda)$, if N is discrete with values in $\mathbb{N}$ such that, for $k \in \mathbb{N}$,

$$\mathbb{P}(N = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

The parameter λ represents the (constant) mean rate of occurence. The Poisson distribution satisfies the following properties.

Proposition 1.2.
Let $N \sim \mathcal{P}(\lambda)$ with $\lambda > 0$.

1. *The Poisson distribution is characterized by its moment-generating function (m.g.f.), or Laplace transform:*
$$\forall\, t \leq 0, \quad \mathbb{E}\left[e^{tN}\right] = \exp\left(\lambda\left(e^t - 1\right)\right).$$
2. *The Poisson distribution is also characterized by its characteristic function:*
$$\forall\, t \in \mathbb{R}, \quad \mathbb{E}\left[e^{itN}\right] = \exp\left(\lambda\left(e^{it} - 1\right)\right).$$

Proof. In exercise. Let $N \sim \mathcal{P}(\lambda)$, with $\lambda > 0$. Let $t \leq 0$.

1. Compute $\mathbb{E}\left[e^{tN}\right]$. ↺

Solution.

Le m.g.f. of N equals

$$\mathbb{E}\left[e^{tN}\right] = \sum_{k\in\mathbb{N}} \left[e^{tk}\left(\frac{\lambda^k}{k!}e^{-\lambda}\right)\right] = \sum_{k\in\mathbb{N}} \left[\frac{(\lambda e^t)^k}{k!}\right] e^{-\lambda} = \exp\left(\lambda e^t - \lambda\right),$$

which ends the proof.

2. Same computation as in 1.

□

Note that both the moment-generating function and the characteristic function characterize the distribution of a random variable. Hence, if two random variables have the same moment-generating function, or the same characteristic function, then they have the same distribution.

Property 1.3.

Let $N \sim \mathcal{P}(\lambda)$ with $\lambda > 0$. Then $\mathbb{E}[N] = \mathrm{Var}(N) = \lambda$.

Proof. In exercise. Recall that if $L : t \mapsto \mathbb{E}\left[e^{tN}\right]$ denotes the moment-generating function, then, for all k in $\mathbb{N}^*$, the kth-order moment of N is linked to the kth-order derivative of L as follows:

$$\mathbb{E}\left[N^k\right] = L^{(k)}(0).$$

Deduce $\mathbb{E}[N]$ and $\mathrm{Var}(N)$.

Solution.

- Expectation: Since for all $t \leq 0$,

$$L'(t) = \lambda e^t \exp\left(\lambda\left(e^t - 1\right)\right) = \lambda \exp\left(\lambda e^t + t - \lambda\right),$$

we directly obtain

$$\mathbb{E}[N] = L'(0) = \lambda.$$

- Variance: Since for all $t \leq 0$,

$$L''(t) = \lambda(\lambda e^t + 1)\exp\left(\lambda e^t + t - \lambda\right),$$

we directly obtain

$$\mathbb{E}\left[N^2\right] = L''(0) = \lambda(\lambda+1),$$

and thus

$$\mathrm{Var}(N) = \mathbb{E}\left[N^2\right] - \mathbb{E}[N]^2 = \lambda.$$

□

II Exponential random variable

The exponential distribution is a continuous probability distribution commonly used to model lifetimes, or more generally time elapsed between events.

Definition 1.4.
Let X be a real random variable. Then X is said to have an exponential distribution with parameter $\lambda > 0$, denoted $X \sim \mathcal{E}(\lambda)$, if X has a density w.r.t. the Lebesgue measure defined on $\mathbb{R}$ by

$$x \mapsto \lambda e^{-\lambda x} \mathbb{1}_{\{x>0\}}.$$

The exponential distribution can be characterized by the following functions.

Proposition 1.5.
Let $X \sim \mathcal{E}(\lambda)$ with $\lambda > 0$.

1. *Its moment-generating function (m.g.f.), or Laplace transform, equals for all $t < \lambda$,*

$$\mathbb{E}\left[e^{tX}\right] = \frac{\lambda}{\lambda - t}.$$

2. *Its distribution functions equals:* $F : t \in \mathbb{R} \mapsto \left(1 - e^{-\lambda t}\right) \mathbb{1}_{\{t \geq 0\}}$.
3. *Its survival function satisfies:* $\forall t \geq 0, \quad \mathbb{P}(X > t) = e^{-\lambda t}$.

Proof. In exercise.

1. a) Let $t < \lambda$. Prove that $\mathbb{E}\left[e^{tX}\right] = \lambda \int_0^{+\infty} e^{-(\lambda - t)x} dx$.
   b) Deduce that $\mathbb{E}\left[e^{tX}\right]$ is well-defined only if $\lambda - t > 0$ and compute the Laplace transform.

   Solution.

$$\mathbb{E}\left[e^{tX}\right] = \int_0^{+\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{+\infty} e^{-(\lambda - t)x} dx \stackrel{(*)}{=} \lambda \left[\frac{-1}{\lambda - t} e^{-(\lambda - t)x}\right]_0^{+\infty} = \frac{\lambda}{\lambda - t}.$$

   Note that the integral $^{(*)}$ is well-defined only if $\lambda - t > 0$.

   Note that in this case, the Laplace transform is positive as expected.

2. Compute $\mathbb{P}(X \leq t)$ by distinguishing the cases $t < 0$ and $t \geq 0$.

   Solution.

   - Since X takes positive values, if $t < 0$, then $\mathbb{P}(X \leq t) = 0$.
   - Let $t \geq 0$. Then

$$\mathbb{P}(X \leq t) = \int_0^t \lambda e^{-\lambda s} ds = \left[-e^{-\lambda s}\right]_0^t = 1 - e^{-\lambda t}.$$

3. Immediate from 2 since $\mathbb{P}(X > t) = 1 - \mathbb{P}(X \leq t)$.

∎

Property 1.6.
Let $X \sim \mathcal{E}(\lambda)$ with $\lambda > 0$. Then $\mathbb{E}[X] = 1/\lambda$ and $\mathrm{Var}(X) = 1/\lambda^2$.

Proof. In exercice, using the moment-generating function (m.g.f.) as in the proof of Property 1.3.

Solution.

Let $X \sim \mathcal{E}(\lambda)$ and denote $L : t \mapsto \mathbb{E}[tX] = \lambda/(\lambda - t)$ its m.g.f.
Then L is differentiable on $]-\infty; \lambda[$, with

$$L'(t) = \frac{\lambda}{(\lambda - t)^2} \qquad \text{and} \qquad L''(t) = \frac{2\lambda}{(\lambda - t)^3}.$$

Hence

$$\mathbb{E}[X] = L'(0) = \frac{1}{\lambda}, \quad \mathbb{E}\left[X^2\right] = L''(0) = \frac{2}{\lambda^2}, \quad \text{and} \quad \mathrm{Var}(X) = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}.$$

□

Property 1.7.

Let $X_1 \sim \mathcal{E}(\lambda_1)$ and $X_2 \sim \mathcal{E}(\lambda_2)$ be two independent exponential random variables. Then

$$\mathbb{P}(X_1 < X_2) = \frac{\lambda_1}{\lambda_1 + \lambda_2}.$$

Proof. In exercise. *Hint:* $\mathbb{P}(X_1 < X_2) = \mathbb{E}\left[\mathbb{1}_{\{X_1 < X_2\}}\right]$.

Solution.

$$\begin{aligned}
\mathbb{P}(X_1 < X_2) &= \mathbb{E}\left[\mathbb{1}_{\{X_1 < X_2\}}\right] \\
&= \int_{\mathbb{R}^2} \mathbb{1}_{\{x_1 < x_2\}} \left(\lambda_1 e^{-\lambda_1 x_1} \mathbb{1}_{\{x_1 > 0\}}\right) \left(\lambda_2 e^{-\lambda_2 x_2} \mathbb{1}_{\{x_2 > 0\}}\right) dx_1 dx_2 \\
&= \int_0^{+\infty} \lambda_1 e^{-\lambda_1 x_1} \left(\int_{x_1}^{+\infty} \lambda_2 e^{-\lambda_2 x_2} dx_2\right) dx_1 \\
&= \int_0^{+\infty} \lambda_1 e^{-\lambda_1 x_1} \left[-e^{-\lambda_2 x_2}\right]_{x_1}^{+\infty} dx_1 \\
&= \int_0^{+\infty} \lambda_1 e^{-(\lambda_1 + \lambda_2) x_1} dx_1 \\
&= \left[\frac{-\lambda_1}{\lambda_1 + \lambda_2} e^{-(\lambda_1 + \lambda_2) x_1}\right]_0^{+\infty} \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}.
\end{aligned}$$

□

Example: A system consists of two independent components in series (the system fails if at least one of the components fails). Denote X_i the time (in years) to failure of component i and assume that $X_1 \sim \mathcal{E}(2)$ and $X_2 \sim \mathcal{E}(6)$. In average, component 1 runs 6 months before failure, whereas component 2 runs only 2 months before failure. So in average, component 2 fails before component 1. In fact, according to Property 1.7, the probability that the first component fails before the second one equals 1/4 (it is coherent).

Proposition 1.8.

Let $(X_i)_{1\leq i\leq n}$ be independent exponential r.v. such that for all $1 \leq i \leq n$, $X_i \sim \mathcal{E}(\lambda_i)$. Then

$$\min_{1\leq i\leq n} X_i \sim \mathcal{E}\left(\sum_{i=1}^{n} \lambda_i\right).$$

Proof. In exercise. *Hint:* Compute the survival function of $\min_{1\leq i\leq n} X_i$.

Solution.

Let $t > 0$. The survival function equals

$$\begin{aligned} \mathbb{P}\left(\min_{1\leq i\leq n} X_i > t\right) &= \mathbb{P}\left(\bigcap_{1\leq i\leq n} \{X_i > t\}\right) \overset{(*)}{=} \prod_{1\leq i\leq n} \mathbb{P}(X_i > t) \\ &= \prod_{1\leq i\leq n} e^{-\lambda_i t} = \exp\left(-\left[\sum_{i=1}^{n} \lambda_i\right] t\right), \end{aligned}$$

where $^{(*)}$ comes from the independence between the X_i's. We recognize the survival function of an exponential r.v. with parameter $(\sum_{i=1}^{n} \lambda_i)$.

□

Example: A system consists of 12 independent components in series with exponential time to failure (in years) $\mathcal{E}(2)$. Individually, the average time to failure of each component is 6 months. However, the time to failure of the system has an exponential distribution $\mathcal{E}(24)$. In particular, its average time to failure is about 15 days.

III Memoryless distribution

Definition 1.9.

A random variable X is said to have a *memoryless distribution* if for all $t, s > 0$,

$$\mathbb{P}(X > t + s | X > t) = \mathbb{P}(X > s)\,. \tag{1.1}$$

Note that this definition is well defined if for all $t > 0$, $\mathbb{P}(X > t) > 0$.

Example: In reliability theory, if we think of X being the time to failure of a component, then (1.1) states that the probability that the component still runs after $s + t$ hours knowing that it is still working after t hours is the same as the initial probability that it runs for at least s hours. In other words, if the component is still working at time t, then the distribution of the remaining amount of time that it works is the same as the original operating distribution; that is, the instrument does not "remember" that is has already been used for a time t.

Proposition 1.10.

The exponential distribution is the only continuous memoryless distribution, i.e.,

1. *If $X \sim \mathcal{E}(\lambda)$, then X has a memoryless distribution.*
2. *If X has a continuous and memoryless distribution, then there exists $\lambda > 0$ such that $X \sim \mathcal{E}(\lambda)$.*

The proof of Proposition 1.10 relies on the following Lemma.

Lemma 1.11.

The functional equation

$$f(x+y) = f(x)f(y) \quad \forall x, y \in \mathbb{R}$$

has a unique continuous solution that is $f : x \in \mathbb{R} \mapsto [f(1)]^x$.

Proof of Proposition 1.10. In exercise.

1. Let $X \sim \mathcal{E}(\lambda)$. Let $t, s > 0$ and compute $\mathbb{P}(X > t+s | X > t)$.

   Solution.

   $$\mathbb{P}(X > t+s|X > t) = \frac{\mathbb{P}(X > t+s)}{\mathbb{P}(X > t)} = \frac{\exp(-\lambda(t+s))}{\exp(-\lambda t)} = \exp(-\lambda s) = \mathbb{P}(X > s)\,.$$

2. Let X be a continuous r.v. with a memoryless distribution. Denote $G : t \mapsto \mathbb{P}(X > t)$ its survival function.

   a) Justify that G is a continuous function.

   Solution.
   Since X is a continuous r.v., then its cumulative distribution function, and thus its survival function, are continuous functions.

   b) Let $s, t > 0$. Prove that $G(t+s) = G(s)G(t)$.

   Solution.
   By definition of the memoryless property,

   $$\begin{aligned} G(t+s) = \mathbb{P}(X > t+s) &= \mathbb{P}(\{X > t+s\} \cap \{X > t\}) \\ &= \mathbb{P}(X > t+s|X > t)\,\mathbb{P}(X > t) \\ &= \mathbb{P}(X > s)\,\mathbb{P}(X > t) = G(s)G(t). \end{aligned}$$

   c) Deduce from Lemma 1.11, that for all $t > 0$, $G(t) = e^{-\lambda t}$ where λ is a well-chosen positive constant (to determine).

   Solution.
   Since G is a continuous function satisfying the functional equality of Lemma 1.11, we can deduce that for all t,

   $$G(t) = [G(1)]^t = \exp(\ln[G(1)]t) = \exp(-\lambda t)\,,$$

   where $\lambda = -\ln[G(1)] > 0$ since $G(1) \in \,]0, 1[$.

   d) Conclude.

   Solution.
   We recognize the survival function of an exponential distribution $\mathcal{E}(\lambda)$.

□

Comment 1.12. Discrete memoryless distributions are geometric distributions.

IV Hazard rate in lifetime models

In this section, we only consider non-negative continuous real-valued random variables, that represent lifetimes. For instance, it may be the time before death or relapse of a patient in a medical study, the time before failure of a component in reliability theory, the time before a new claim in actuarial science, or the time before an aircraft loses performance in aeronautics.

The hazard rate represents the instantaneous probability of occurence of the event. It is mathematically defined as follows.

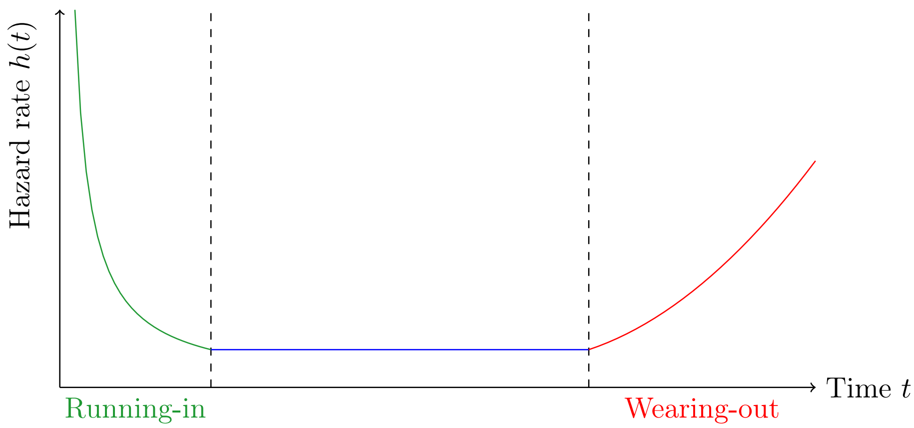
Definition 1.13.
Let X be a non-negative real random variable with density function f w.r.t. the Lebesgue measure. The hazard rate function h is defined by

$$h(t) = \begin{cases} \dfrac{f(t)}{\mathbb{P}(X > t)} & \text{if } \mathbb{P}(X > t) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Remark: Notice that, heuristically,

$$h(t)dt = \frac{f(t)dt}{\mathbb{P}(X > t)} \approx \frac{\mathbb{P}(X \in \,]t, t + dt])}{\mathbb{P}(X > t)} = \mathbb{P}(X \in \,]t + dt] \, | X > t) \, .$$

For instance, in reliability theory, when X denotes the (random) time to failure of a component, the hazard rate corresponds to the probability of a failure occurring juste after time t given that the device is still working at time t. The survival function $R : t \mapsto \mathbb{P}(X > t)$ is called *reliability function*, and the hazard rate is called *failure rate* and is usually denoted λ. In particular, if the failure rate h is decreasing, it means that the instantaneous probability of failure decreases with the age; the component state improves. On the contrary, if the failure rate h is increasing, it means that the instantaneous probability of failure increases with the age; the component deteriorates. In the reliability theory, the failure rate is usually shaped as a bathtub curve (see below).



The survival function and the hazard rate are linked as below.

Proposition 1.14.
Let X be a continuous non-negative r.v. The hazard rate of X is h, if and only if, for all $t > 0$,

$$\mathbb{P}(X > t) = \exp\left(- \int_0^t h(x) dx \right) .$$

In particular, if X and Y are two continuous r.v with the same hazard rate, then they have the same distribution.

Proof. In exercise. Denote $G : t \mapsto \mathbb{P}(X > t)$ the survival function of X.

$\boxed{\Rightarrow}$ Denote h the hazard rate of X.

a) Prove that for all $t > 0$, $\mathbb{P}(X \leq t) = \int_0^t h(s)\mathbb{P}(X > s)\, ds$, and deduce that

$$G(t) = 1 - \int_0^t h(s)G(s)ds. \tag{1.2}$$

Solution.
Since the hazard rate of X is h, then its density equals for all $s > 0$, $f(s) = h(s)\mathbb{P}(X > s)$. In particular, its c.d.f equals

$$\mathbb{P}(X \leq t) = \int_0^t f(s)ds = \int_0^t h(s)\mathbb{P}(X > s)\, ds,$$

that is $1 - G(t) = \int_0^t h(s)G(s)ds$.

b) Deduce that G is differentiable and satisfies

$$\begin{cases} G(0) = g_0 \\ G'(t) = -h(t)G(t), \end{cases}$$

where g_0 is a constant to determine.

Solution.
From (1.2), we directly deduce that G is differentiable and satisfies $-G'(t) = h(t)G(t)$. Moreover, $G(0) = \mathbb{P}(X > 0) = 1$.

c) Deduce the expression of $G(t)$.

Solution.
G is the unique solution of the homogeneous linear differential equation with initial condition $G(0) = 1$. Hence,

$$G(t) = \underbrace{G(0)}_{1} \exp\left(-\int_0^t h(s)ds\right).$$

$\boxed{\Leftarrow}$ Assume that there exists a function h s.t. the survival function $G(t) = \mathbb{P}(X > t) = \exp\left(-\int_0^t h(x)dx\right)$.

i) Compute the cumulative distribution function of X and prove that its density function is $f : t \mapsto f(t) = h(t)G(t)$.

Solution.
The c.d.f. of X equals for all $t > 0$

$$F(t) = 1 - G(t) = 1 - \exp\left(-\int_0^t h(x)dx\right).$$

Hence, the density of X equals for all $t > 0$

$$f(t) = F'(t) = -\left[-h(t)\ \exp\left(-\int_0^t h(x)dx\right)\right] = h(t)G(t).$$

ii) Deduce the hazard rate of X.

Solution.

Hence, by definition, the hazard rate of X equals

$$\frac{f(t)}{G(t)} = h(t),$$

which ends the proof.

□

Property 1.15.

Let $\lambda > 0$. The random variable $X \sim \mathcal{E}(\lambda)$ if and only if its hazard rate function is constant: $h(t) = \lambda$ for all $t > 0$.

Proof. In exercise.

$\boxed{\Rightarrow}$ Assume $X \sim \mathcal{E}(\lambda)$ and compute the hazard rate function of X.

Solution.

Recall the density function f of X equals for all $t > 0$, $f(t) = \lambda e^{-\lambda t}$, and the survival function is $\mathbb{P}(X > t) = e^{-\lambda t}$. Hence the hazard rate function of X equals

$$h(t) = \frac{f(t)}{\mathbb{P}(X > t)} = \lambda.$$

$\boxed{\Leftarrow}$ Assume the hazard rate function of X is constant equal to λ. Prove that $X \sim \mathcal{E}(\lambda)$ using Proposition 1.14 without any computation.

Solution.

By Proposition 1.14, the hazard rate characterizes the distribution. Hence a r.v. with constant hazard rate function equal to λ has the same hazard rate function, and thus the same distribution than an exponential r.v. (by $\boxed{\Rightarrow}$), that is $\mathcal{E}(\lambda)$.

□

V On the gamma and Erlang's distributions

The Erlang distribution generalizes the exponential distribution. It is commonly used to model the waiting time between more than one events. It was developed by A. K. Erlang to examine the number of telephone calls which might be made at the same time to the operators of the switching stations. More precisely, the Erlang distribution is the distribution of the sum of n i.i.d. r.v. with exponential distribution. The gamma distribution then generalizes the Erlang distribution by allowing n to be any positive real number.

Definition 1.16.

- **Gamma distribution.** Let $\lambda > 0$ and $\alpha > 0$. A non-negative real-valued random variable

X with density (w.r.t. the Lebesgue measure)

$$x \mapsto \frac{(\lambda x)^{\alpha-1}}{\Gamma(\alpha)} \lambda e^{-\lambda x} \mathbb{1}_{\{x>0\}},$$

where $\Gamma(\alpha) = \int_0^{+\infty} e^{-t} t^{\alpha-1} dt$ is the gamma function, is said to have a gamma distribution with parameter (α, λ). We denote $X \sim \Gamma(\alpha, \lambda)$.

- **Erlang distribution.** If $\alpha = n \in \mathbb{N}^*$, then the gamma distribution $\Gamma(n, \lambda)$ is called Erlang's distribution. In particular, since $\Gamma(n) = (n-1)!$, the density of a r.v. $X \sim \Gamma(n, \lambda)$ equals

$$x \mapsto \frac{\lambda^n x^{n-1}}{(n-1)!} e^{-\lambda x} \mathbb{1}_{\{x>0\}}.$$

The gamma distribution can be characterized by its moment-generating function.

Proposition 1.17.

Let $\alpha, \lambda > 0$ and consider $X \sim \Gamma(\alpha, \lambda)$.

1. *The moment-generating function, or Laplace transform of X equals for all $t < \lambda$,*

$$\mathbb{E}\left[e^{tX}\right] = \left(\frac{\lambda}{\lambda - t}\right)^{\alpha}.$$

2. *We deduce that $\mathbb{E}[X] = \alpha/\lambda$ and $\mathrm{Var}(X) = \alpha/\lambda^2$.*

Proof. In exercise (Worksheet 1, Exercise 2). □

Proposition 1.18.

Links with the exponential and the Chi-square distributions.

i) If $X \sim \Gamma(\alpha, \lambda)$, then, for all $\mu > 0$, $\mu X \sim \Gamma(\alpha, \lambda/\mu)$.

ii) $\Gamma(1, \lambda) \sim \mathcal{E}(\lambda)$.

iii) If $X_1, \dots, X_n$ are i.i.d. with distribution $\mathcal{E}(\lambda)$, then $\sum_{i=1}^{n} X_i \sim \Gamma(n, \lambda)$.

iv) $\Gamma(n, 1/2) \overset{(d)}{=} \chi^2(2n)$.

Proof. In exercise.

i) Use the m.g.f. to prove this point.

Solution.

For all $t > 0$, the m.g.f. of μX equals

$$\mathbb{E}\left[e^{t(\mu X)}\right] = \mathbb{E}\left[e^{(\mu t)X}\right] = \left(\frac{\lambda}{\lambda - \mu t}\right)^{\alpha} = \left(\frac{\lambda/\mu}{\lambda\mu - t}\right)^{\alpha}.$$

We recognize the m.g.f. of a $\Gamma(\alpha, \lambda/\mu)$ distribution.

ii) By taking $\alpha = 1$ in the definition, since $\Gamma(1) = 0! = 1$, we recognize the density of an exponential r.v. with parameter λ.

iii) Worksheet 1, Exercise 2.

iv) Immediate since the density of a $\chi^2(d)$ is

$$x \mapsto \frac{1}{2^{d/2}\Gamma(d/2)} x^{d/2-1} e^{-x/2} \mathbb{1}_{\{x>0\}}.$$

$\square$

Chapter 2

Homogeneous Poisson processes

Homogeneous Poisson processes may provide very simple models in Reliability theory. For instance, if at each failure a system is repaired as new and placed in service again, then the failures can be modeled by an homogeneous Poisson process, provided the repair times can be neglected.

I Definitions

Before defining homogeneous Poisson processes, let us now introduce the notion of point processes on $\mathbb{R}_+$. We observe the occurrences times of a phenomenon, such as failures in reliability theory, or claims in actuarial sciences.

I.1 Point processes and counting processes

There are several ways of defining a point process depending which quantity is considered.

A) Counting process

The first perspective consists in counting the number of events occurring in any time interval (or time window).

Definition 2.1 (Counting process).

A stochastic process $N = (N_t)_{t \in \mathbb{R}_+}$ is said to be a *counting process* if, for all $t \geq 0$, N_t represents the total number of "events" that occur by time t.

Example 2.2.

Here are some examples of counting processes.

a) In reliability theory, one may denote N_t the number of failures of a reparable component or before time t. An event thus corresponds to a failure.

b) In actuarial science, N_t may represent the number of accidents an insurance company has to cover by time t. Here an event is a claim.

c) If we say an event occurs whenever a child is born, then N_t equals the total number of people who where born by time t.

d) Let N_t count the number of goals scored by the French team during the Football World Cup final after t minutes of game. An event if this process occurs when a player scores a goal.

Property 2.3.

A counting process satisfies the following properties.

(i) For all $t \geq 0$, $N_t \geq 0$.

(ii) $t \mapsto N_t$ is non-decreasing, that is for all $s \leq t$, $N_s \leq N_t$.

(iii) $t \mapsto N_t$ is piecewise constant, with values in $\mathbb{N}$.

(iv) $(N_t)_{t \in \mathbb{R}_+}$ is càdlàg, that is right-continuous and left-limited (for *continue à droite, limitée à gauche* in French).

(v) For all $0 \leq s < t$, $N_t - N_s$ counts the number of events that occur in the interval $(s, t]$.

In this lecture, we restrict attention to counting processes N whose realizations are *locally finite* (i.e. the number of events in any bounded domain in $\mathbb{R}_+$ is finite). In particular, this is true as soon as for all $t \geq 0$, N_t is almost surely finite.

B) Point process

The second perspective consists in considering the times each "event" occurs.

Definition 2.4 (Point process).

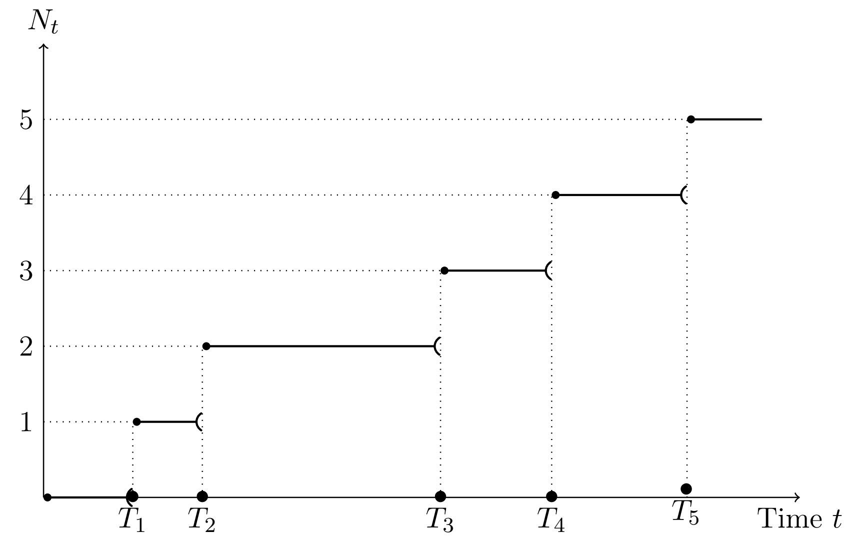
A point process on $\mathbb{R}_+$ is a random countable subset of $\mathbb{R}_+$, each point representing the time occurence of an "event". It may be formally defined as the intersection between $\mathbb{R}_+$ and the non-decreasing sequence of times the event occurs, also called *arrival times*, and denoted

$$0 \leq T_1 \leq T_2 \leq \ldots \leq T_n \leq \ldots$$

where the T_n belong to $\mathbb{R}_+ \cup \{+\infty\}$.

The knowledge of the counting process N is equivalent to the knowledge of the arrival times

$$0 \leq T_1 \leq T_2 \leq \ldots \leq T_n \leq \ldots$$



Indeed, on the one hand, if we know the counting process, it is possible to recover the arrival times as

$$T_n = \inf\{t \geq 0 \; ; \; N_t \geq n\}, \tag{2.1}$$

with the convention that $\inf\{\varnothing\} = +\infty$. On the other hand, if we know the point process, it is possible to define the corresponding counting process for all $t \geq 0$ by

$$N_t = \sum_{n \geq 1} \mathbb{1}_{\{T_n \leq t\}}. \tag{2.2}$$

Notice that this sum converges as soon as the point process does not have any accumulation point, that is the process is locally finite.

Proposition 2.5 (Link between counting and point processes).

Let N be a counting process and denote $(T_n)_{n\in\mathbb{N}^*}$ the corresponding arrival times. Then, for all n in $\mathbb{N}^*$ and all $t \geq 0$,

(i) $\{T_n \leq t\} = \{N_t \geq n\}$.

(ii) $\{T_n > t\} = \{N_t < n\}$.

(iii) $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$.

More generally, for all n in $\mathbb{N}^*$ and all $(t_1, t_2, \ldots, t_n)$ in $(\mathbb{R}_+)^n$,

$$\{T_1 \leq t_1, T_2 \leq t_2, \ldots, T_n \leq t_n\} = \{N_{t_1} \geq 1, N_{t_2} \geq 2, \ldots, N_{t_n} \geq n\}.$$

It is thus equivalent to define the distribution of the point process by the one of $(T_n)_{n\geq 1}$ or the one of $(N_t)_{t\in\mathbb{R}_+}$.

From now on, unless specified, N refers to both the counting process $(N_t)_{t\in\mathbb{R}_+}$ and its associated point process $(T_n)_{n\in\mathbb{N}^*}$, and we do not distinguish between "counting" and "point" processes.

C) Some basic concepts

In this lecture, we only consider *simple* point processes, for which several events cannot occur at the same time.

- On the one hand, from a "point process" perspective, this means that the random variables (T_n) are a.s. pairwise distinct on $\mathbb{R}_+$. More precisely, for all n in $\mathbb{N}^*$ and almost all ω in Ω,

  $$T_n(\omega) < +\infty \quad \Longrightarrow \quad T_n(\omega) < T_{n+1}(\omega).$$

  Essentially, it means that for almost all $\omega \in \Omega$, the sequence $(T_n(\omega))_{n\in\mathbb{N}^*}$ is strictly increasing as long as the $T_n(\omega)$ are finite.
- On the other hand, from a "counting process" perspective, it means that all the jumps of the counting process $(N_t)_{t\in\mathbb{R}_+}$ are equal to 1.

Definition 2.6 (Regular point process).

A point process N is said to be *regular* if for all $t \geq 0$,

$$\lim_{h\to 0} \frac{\mathbb{P}(N_{t+h} - N_t \geq 2)}{h} = 0.$$

For instance in reliability theory, it means that the system will not experience two or more failures simultaneously. In particular, a regular point process is simple.

Definition 2.7 (Mean function and rate of a counting process).

Let N be a counting process.

- The **mean function** m of N is defined by

$$m : t \in \mathbb{R}_+ \mapsto m(t) = \mathbb{E}[N_t] \, .$$

- The **rate** w of N is defined by

$$w : t \in \mathbb{R}_+ \mapsto w(t) = m'(t) = \lim_{h \to 0} \frac{\mathbb{E}[N_{t+h} - N_t]}{h} .$$

Comment 2.8.

i. Note that the mean function and the rate of a counting process are deterministic.

ii. Moreover, even if the counting process is piecewise constant (it has "jumps"), its mean function is not necessarily discontinuous.

iii. If N is a regular point process, the probability of two or more events in $]t, t+h]$ is negligible when h is small. Thus, we may assume that, for small h, $N_{t+h} - N_t$ takes values in $\{0, 1\}$, and that

$$\mathbb{E}[N_{t+h} - N_t] \approx \mathbb{P}(N_{t+h} - N_t = 1) \, .$$

Hence, rate of the process is approximately equal to

$$w(t) \approx \frac{\mathbb{P}(N_{t+h} - N_t = 1)}{h} .$$

Hence, we can think of $w(t)dt$ as the probability of an occurence "just after time t".

iv. In reliability theory, the rate is called **Rate of Occurence of Failure** (RoCoF). It represents the mean number of failures per unit of time.

v. The rate can be generalized to the *conditional rate*, in the case where we work conditionally on all the history of the process. Unlike the rate, the conditional rate is usually stochastic. In reliability theory, it is called **Failure intensity**.

D) Increments of a point process

Definition 2.9 (Stationary increments).

A point process N is said to have stationary increments if the distribution of the number of points in any interval of time depends only on the length of the interval. In other words, for any $s \geq 0$, the number of points in the interval $(t, t+s]$, that is $N_{t+s} - N_t$ has the same distribution for all t.

Example (2.2 continued).

Recall the examples enumerated above.

a) In reliability theory, the stationarity seams reasonable if we consider that there are no running-in period or aging effect, and if after a repair, the component is "as new" (this case is often referred to as **perfect repair**).

b) In actuarial science, it is a reasonable assumption if, for instance, there are no periods in the year where the accidents seem more likely to happen.

c) If we believe that the earth's population is basically constant (a belief not held at present by most scientists), then the assumption of stationary increments might be reasonable.

d) Not reasonable if we take into account the tiredness of the players.

Definition 2.10 (Independent increments**).**
A point process N is said to have independent increments if the number of points belonging to disjoint intervals are independent. In other words, for all n in $\mathbb{N}^*$, for all $0 < t_1 < t_2 < \ldots < t_n$, the random variables $N_{t_1} - N_0$, $N_{t_2} - N_{t_1}$, $\ldots$, $N_{t_n} - N_{t_{n-1}}$ are independent.

Example (2.2 continued).
Recall the examples enumerated above.

a) In reliability theory, it seams reasonable to assume that the increments are independent if we can assume that a failure does not have repercussions (e.g. **perfect repair** case).

b) It depends on the kind of insurance: in life insurance, the increments can be assumed as independent (the number of deaths at time t does not impact the other customers). Yet, in non-life insurance, the independence assumption for the increments is less realistic. For instance, if an earthquake happens, then there will probably be aftershocks and thus other damages.

c) No. If, for instance, N_t is very large, then it is probable that there are many people alive at time t; this would lead us to believe that the number of new births between time t and time $t + s$ would also tend to be large. Hence it does not seem reasonable to assume that N_t is independent of $N_{t+s} - N_t$.

d) Could be justified if we assume the players' chances of scoring a goal do not depend on how the game has been going (unlikely).

I.2 Homogeneous Poisson process on $\mathbb{R}_+$

The most elementary point process is the homogeneous Poisson point process which is defined as follows.

Definition 2.11.
A point process N is said to be a homogeneous Poisson process with rate $\lambda > 0$ if

(1) $N_0 = 0$.

(2) The process has independent increments.

(3) The number of points in any interval of length $t > 0$ has a Poisson distribution with parameter λt, that is, for all $s \geq 0$,

$$N_{s+t} - N_s \sim \mathcal{P}(\lambda t).$$

The parameter λ is also called the intensity of the Poisson point process.

Property 2.12.
Let N be a homogeneous Poisson process with intensity $\lambda > 0$.

1. N has stationary increments.

2. The mean function of N equals

$$\forall t \geq 0, \quad m(t) = \mathbb{E}[N_t] = \lambda t.$$

   In particular, the rate (see Definition 2.7) of the Poisson process N is constant (equal, as expected, to λ):

$$\forall t \geq 0, \quad w(t) = \lambda.$$

3. One has as $t \to 0$,

   (a) $\mathbb{P}(N_t = 0) = e^{-\lambda t} = 1 - \lambda t + o(t)$.

   (b) $\mathbb{P}(N_t = 1) = \lambda t e^{-\lambda t} = \lambda t + o(t)$.

   (c) $\mathbb{P}(N_t \geq 2) = \sum_{k=2}^{+\infty} \frac{(\lambda t)^k}{k!} e^{-\lambda t} = o(t)$.

Proof. Immediate by Definition 2.11, and since $N_t \sim \mathcal{P}(\lambda t)$. $\square$

Comment 2.13. In particular, $\forall t \geq 0$,

$$\frac{\mathbb{P}(N_{t+h} - N_t \geq 2)}{h} = \frac{\mathbb{P}(N_h \geq 2)}{h} = o(1) \underset{h \to 0}{\longrightarrow} 0.$$

Hence a homogeneous Poisson process is regular (and thus simple).

Definition 2.14.
The point process N is said to be a homogeneous Poisson process with rate (or with intensity) $\lambda > 0$ if

(A) $N_0 = 0$.

(B) The process has stationary and independent increments.

(C) As $t \to 0$, $t > 0$,

   (a) $\mathbb{P}(N_t = 1) = \lambda t + o(t)$.

   (b) $\mathbb{P}(N_t \geq 2) = o(t)$.

Proposition 2.15.
Definition 2.11 and Definition 2.14 are equivalent.

Proof. In exercise*.

$\boxed{\Rightarrow}$ Done in Property 2.12.

$\boxed{\Leftarrow}$ Assume Definition 2.14. Then, the points (1) and (2) in Definition 2.11 are immediate from (A) and (B) in Definition 2.14.
Remains to prove (3). To do so, let us compute the moment-generating function of N_t for all $t > 0$.

a) Let $h > 0$. Using Assumption (C), prove that for all $u \leq 0$,

$$\mathbb{E}\big[e^{uN_h}\big] = 1 + \lambda h(e^u - 1) + o(h).$$

Solution.

From (C), we obtain

$$\mathbb{P}(N_h = 0) = 1 - [\mathbb{P}(N_h = 1) + \mathbb{P}(N_h \geq 2)] = 1 - \lambda h + o(h).$$

Hence, since $u \leq 0$,

$$\begin{aligned} \mathbb{E}\left[e^{uN_h}\right] &= \underbrace{\mathbb{P}(N_h = 0)}_{1-\lambda h+o(h)} + e^u \underbrace{\mathbb{P}(N_h = 1)}_{\lambda h+o(h)} + \underbrace{\sum_{k\geq 2} e^{uk}\mathbb{P}(N_h = k)}_{\leq \sum_{k\geq 2}\mathbb{P}(N_h=k)=\mathbb{P}(N_h\geq 2)=o(h)} \\ &= 1 + \lambda h(e^u - 1) + o(h). \end{aligned}$$

b) Fix $u \leq 0$ and denote

$$g : t \mapsto \mathbb{E}\left[e^{uN_t}\right].$$

We now aim at proving that for any small h,

$$g(t+h) = g(t)\left[1 + \lambda h(e^u - 1) + o(h)\right]. \tag{2.3}$$

i) Let $h > 0$. Using both properties in Assumption (B), prove that

$$g(t+h) = \mathbb{E}\left[e^{uN_h}\right] g(t).$$

and deduce (2.3). *Hint:* $N_{t+h} = (N_{t+h} - N_t) + N_t$.

Solution.

By stationarity of the increments,

$$(N_{t+h} - N_t) \text{ has the same distribution than } N_h \qquad (\star),$$

and by independence of the increments, since $(0;t] \cap (t;t+h] = \varnothing$,

$$(N_{t+h} - N_t) \quad \perp\!\!\!\perp \quad N_t - N_0 = N_t \qquad (\dagger).$$

Therefore,

$$\begin{aligned} g(t+h) = \mathbb{E}\left[e^{uN_{t+h}}\right] &= \mathbb{E}\left[e^{u(N_{t+h}-N_t)}e^{uN_t}\right] \\ &\overset{(\dagger)}{=} \mathbb{E}\left[e^{u(N_{t+h}-N_t)}\right]\mathbb{E}\left[e^{uN_t}\right] \\ &\overset{(\star)}{=} \mathbb{E}\left[e^{uN_h}\right] g(t). \end{aligned}$$

Equation (2.3) is then immediate from a).

ii) Let $h < 0$. In the same way, prove that

$$g(t) = \mathbb{E}\left[e^{uN_{(-h)}}\right] g(t+h),$$

and deduce (2.3).

Solution.

As in the previous case, since $N_t = (N_t - N_{t+h}) + N_{t+h}$, by stationarity and independence (since $(0;t+h]\cap(t+h;t]=\varnothing$) of the increments, we obtain

$$\begin{aligned} g(t)=\mathbb{E}\left[e^{uN_t}\right] &= \mathbb{E}\left[e^{u(N_t-N_{t+h})}e^{uN_{t+h}}\right] \\ &= \mathbb{E}\left[e^{u(N_t-N_{t+h})}\right]\mathbb{E}\left[e^{uN_{t+h}}\right] \qquad \text{(independence)} \\ &= \mathbb{E}\left[e^{uN_{(-h)}}\right]g(t+h) \qquad \text{(stationarity)}. \end{aligned}$$

We deduce from a) that $g(t)=g(t+h)\left[1-\lambda h(e^u-1)+o(h)\right]$, and thus

$$g(t+h)=\frac{g(t)}{\left[1-\lambda h(e^u-1)+o(h)\right]}=g(t)\left[1+\lambda h(e^u-1)+o(h)\right].$$

c) i. Deduce from (2.3) that g is differentiable at point t, and satisfies the homogeneous linear differential equation with initial condition $g(0)$ (to determine)

$$\begin{cases} g(0)=\dots \\ g'(t)=g(t)\lambda(e^u-1). \end{cases} \tag{2.4}$$

Solution.

From (2.3), we directly deduce that

$$\frac{g(t+h)-g(t)}{h}=g(t)\left[\lambda(e^u-1)+o(1)\right]$$

has a finite limit when $h\to 0$, which equals

$$g'(t)=\lim_{h\to 0}\frac{g(t+h)-g(t)}{h}=g(t)\left[\lambda(e^u-1)\right].$$

Moreover, $g(0)=\mathbb{E}\left[e^0\right]=1$.

ii. Solve the differential equation and deduce that

$$g(t)=e^{\lambda t(e^u-1)}.$$

Solution.

Hence g is the unique solution of the homogeneous linear differential equation with initial condition (2.4), that is

$$g(t)=g(0)e^{\lambda(e^u-1)t}=e^{\lambda t(e^u-1)}.$$

d) Deduce that $N_t\sim\mathcal{P}(\lambda t)$, and then (3).

Solution.

We proved that for any $t>0$ and any $u\le 0$,

$$\mathbb{E}\left[e^{uN_t}\right]=e^{\lambda t(e^u-1)}.$$

We recognize the m.g.f. of a $\mathcal{P}(\lambda t)$ distributed r.v. (c.f. Proposition 1.2).

Finally, by the stationarity assumption, for all $s\ge 0$, $N_{s+t}-N_s$ and N_t have the same distribution, that is $\mathcal{P}(\lambda t)$.

□

II Arrival and interarrival times

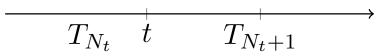
In all this section, N denotes a point process, and $(T_n)_{n\in\mathbb{N}^*}$ is the sequence of the corresponding arrival times (see Equations (2.1) and (2.2) for more details).

II.1 First arrival time after a given instant

Property 2.16.

Let N be a homogeneous Poisson process with rate $\lambda > 0$ and denote $(T_n)_{n\in\mathbb{N}^*}$ the corresponding arrival times.

- Then, the first arrival time satisfies $T_1 \sim \mathcal{E}(\lambda)$.
- Moreover, for all fixed $t > 0$, one has



$$T_{N_t} \leq t < T_{N_t+1}.$$

In particular, the first arrival time after a given instant t satisfies

$$T_{N_t+1} - t \sim \mathcal{E}(\lambda).$$

Proof. In exercise (c.f. Worksheet 2, Exercise 2). □

Example (The inspection paradox or *"le paradoxe de l'autobus"*).

You are waiting for a bus to arrive. It is written on your timetable that there is a bus every 10 minutes.

- If the arrival times of the bus were deterministic, and a bus arrived every 10 minutes, then the time you wait at the bus stop is uniformly distributed on $[0, 10]$ and, in expectation you wait for $10/2 = 5$ minutes.
- However, due to traffic, the arrival times of the buses are random. You assume they can be modeled by a Poisson process with rate $1/10$ (since in average, there is 1 bus in 10 min). Then, your waiting time equals $T_{N_t+1} - t$ which has and exponential distribution with parameter $1/10$. Hence, in average, you wait for 10 minutes (instead of the expected 5 mins).

II.2 Exponential interarrival times

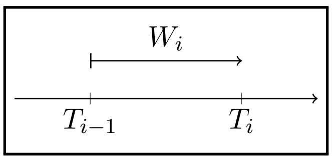
Definition 2.17.

The interarrival times (also called waiting times) of a point process are defined by

$$\begin{cases} W_1 &= T_1, \\ W_i &= T_i - T_{i-1}, \quad \forall i \geq 2. \end{cases}$$

The ith waiting time corresponds to the delay between the $(i-1)$th and the ith occurrences. In particular, for all n in $\mathbb{N}^*$,

$$T_n = \sum_{i=1}^{n} W_i.$$



Proposition 2.18.

Let N is a homogeneous Poisson process with rate λ.

1. *Then the corresponding interarrival times $(W_i)_{i\geq 1}$ are independent and identically distributed exponential random variables with parameter λ.*
2. *It follows that for all n in $\mathbb{N}^*$, the nth arrival time T_n has a gamma (or Erlang) distribution with parameters n and λ.*

Notice that, heuristically, the assumptions of stationarity and independent increments means that, at any point in time, the process *probabilistically* restarts itself, independently on the past. That is, the process from any time t is independent of all that has previously occurred, and also has the same distribution as the original process. In other words, the process has *no memory*, and hence, exponential interarrival times are to be expected (c.f. Section III of Chapter 1).

This result relies on the following Lemma.

Lemma 2.19.

Let N be a homogeneous Poisson process with rate λ, and denote $T_1 < T_2, \ldots < T_n < \ldots$ the corresponding arrival times. Then, for all n in $\mathbb{N}^*$, $(T_1, \ldots, T_n)$ has the following density w.r.t. the Lebesgue measure on $\mathbb{R}^n$:

$$(t_1, \ldots, t_n) \mapsto \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}}.$$

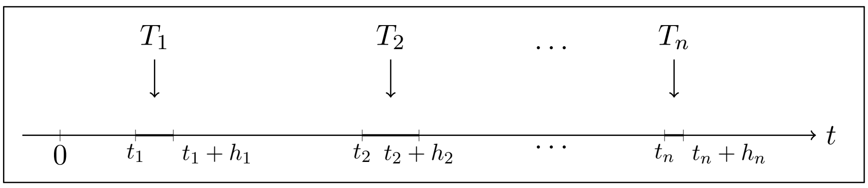
Idea of proof of Lemma 2.19.

- (Reminder) **For a real-valued continuous random variable:** Let X be a random variable with density f and c.d.f. F. Then, its density

$$f(x) = F'(x) = \lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{h\to 0} \frac{\mathbb{P}(X \in (t; t+h])}{h}.$$

- **In the multidimensional case:** Similarly, consider a strictly increasing sequence $0 < t_1 < \ldots < t_n$, and introduce for $h = (h_1, \ldots, h_n)$ small enough such that all $(t_i, t_i + h_i]$ are disjoint,

$$A_h = \{T_1 \in (t_1; t_1 + h_1], \ldots, T_n \in (t_n; t_n + h_n]\}.$$



Then,

$$f(t_1, \ldots, t_n) = \lim_{\forall i,\ h_i \to 0} \frac{\mathbb{P}(A_h)}{h_1 \ldots h_n}.$$

Yet, the event A_h is satisfied if and only if there are

- 0 points in $(0, t_1]$ and 1 point in $(t_1, t_1 + h_1]$,
- 0 points in $(t_1 + h_1, t_2]$ and 1 point in $(t_2, t_2 + h_2]$,
- ...
- 0 points in $(t_{n-1} + h_{n-1}, t_n]$ and 1 point in $(t_n, t_n + h_n]$.

Thus, since for $h_1, \ldots, h_n$ small enough, the intervals appearing in A_h are disjoints, by independence of the increments,

$$\begin{aligned}
\mathbb{P}(A_h) &= \mathbb{P}(N_{t_1} = 0) \times \mathbb{P}(N_{t_1+h_1} - N_{t_1} = 1) \times \mathbb{P}(N_{t_2} - N_{t_1+h_1} = 0) \times \mathbb{P}(N_{t_2+h_2} - N_{t_2} = 1) \times \\
&\quad \times \ldots \times \ \mathbb{P}\big(N_{t_n} - N_{t_{n-1}+h_{n-1}} = 0\big) \times \mathbb{P}(N_{t_n+h_n} - N_{t_n} = 1) \\
&= e^{-\lambda t_1} \times \left[\lambda h_1 e^{-\lambda h_1}\right] \times e^{-\lambda(t_2 - [t_1+h_1])} \times \left[\lambda h_2 e^{-\lambda h_2}\right] \times \ldots \times e^{-\lambda(t_n - [t_{n-1}+h_{n-1}])} \times \left[\lambda h_n e^{-\lambda h_n}\right] \\
&= \lambda^n h_1 \ldots h_n e^{-\lambda(t_n+h_n)}.
\end{aligned}$$

Hence,

$$f(t_1, \ldots, t_n) = \lim_{\forall i,\ h_i \to 0} \lambda^n e^{-\lambda(t_n+h_n)} = \lambda^n e^{-\lambda t_n}.$$

Note that, if we do not have $0 < t_1 < \ldots < t_n$, then $\mathbb{P}(A_h) = 0$ since $0 < T_1 < \ldots < T_n$ a.s.

□

We can now prove Proposition 2.18.

Proof. In exercise.

1. Let N be a Poisson process with rate λ, and consider $(W_i)_{i\geq 1}$ the interarrival times. Let us prove that they are i.i.d. with distribution $\mathcal{E}(\lambda)$.

   (a) Fix $n \in \mathbb{N}^*$. Consider $g : \mathbb{R}^n \to \mathbb{R}^n$ a measurable bounded function. Prove that

$$\mathbb{E}[g(W_1, \ldots, W_n)] = \int_{\mathbb{R}^n} g(t_1, t_2 - t_1, \ldots, t_n - t_{n-1}) \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}} dt_1 \ldots dt_n.$$

Solution.

Consider the arrival times $T_i = \sum_{i=1}^n W_i$. Knowing the denisty of $(T_1, \ldots, T_n)$ given in Lemma 2.19 allows us to write

$$\begin{aligned}
\mathbb{E}[g(W_1, \ldots, W_n)] &= \mathbb{E}[g(T_1, T_2 - T_1, \ldots, T_n - T_{n-1})] \\
&= \int_{\mathbb{R}^n} g(t_1, t_2 - t_1, \ldots, t_n - t_{n-1}) \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}} dt_1 \ldots dt_n
\end{aligned}$$

   (b) By a well-chosen change of variables (see Appendix I), show that

$$\mathbb{E}[g(W_1, \ldots, W_n)] = \int_{\mathbb{R}^n} g(w_1, w_2, \ldots, w_n) \prod_{i=1}^n \left[\lambda e^{-\lambda w_i} \mathbb{1}_{\{w_i>0\}}\right] dw_1 \ldots dw_n.$$

Solution.

Consider the following change of variable:

$$\begin{cases} w_1 = t_1 \\ w_2 = t_2 - t_1 \\ \vdots \\ w_n = t_n - t_{n-1} \end{cases} \quad \Leftrightarrow \quad \begin{cases} t_1 = w_1 & = \phi_1(w) \\ t_2 = w_1 + w_2 & = \phi_2(w) \\ \vdots & \vdots \\ t_n = \sum_{i=1}^n w_n & = \phi_n(w) \end{cases}$$

The function $\phi : w = (w_1, \ldots, w_n) \mapsto (w_1, w_1 + w_2, \ldots, \sum_{i=1}^n w_i)$ is a one-to-one $\mathcal{C}^1$-

diffeomorphism with Jacobian matrix

$$J_\phi(w) = \begin{pmatrix} \frac{\partial \phi_1}{\partial w_1} & \dots & \frac{\partial \phi_1}{\partial w_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial w_1} & \dots & \frac{\partial \phi_n}{\partial w_n} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

Hence, for all w in $\mathbb{R}^n$, $\det(J_\phi(w)) = 1 \neq 0$. Moreover,

$$\mathbb{1}_{\{0<t_1<\ldots<t_n\}} = \mathbb{1}_{\left\{\bigcap_{i=1}^n \{w_i>0\}\right\}} = \prod_{i=1^n} \mathbb{1}_{\{w_i>0\}}.$$

Hence, we obtain the following:

$$\begin{aligned} &\int_{\mathbb{R}^n} g(t_1, t_2 - t_1, \ldots, t_n - t_{n-1})\lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}} dt_1 \ldots dt_n \\ &\qquad = \int_{\mathbb{R}^n} g(w_1, w_2, \ldots, w_n)\lambda^n e^{-\lambda \sum_{i=1}^n w_i} \prod_{i=1^n} \mathbb{1}_{\{w_i>0\}} \left|\det(J_\phi(w))\right| dw_1 \ldots dw_n \\ &\qquad = \int_{\mathbb{R}^n} g(w_1, w_2, \ldots, w_n) \underbrace{\prod_{i=1}^n \left[\lambda e^{-\lambda w_i} \mathbb{1}_{\{w_i>0\}}\right]}_{(*)} dw_1 \ldots dw_n. \end{aligned}$$

(c) Conclude.

Solution.

We recognize in $(*)$ the density of n i.i.d. r.v. with $\mathcal{E}(\lambda)$ distribution.

2. Immediate from Proposition 1.18, iii).

□

The reverse of Proposition 2.18 is true and thus provides a way of constructing a homogeneous Poisson process (see Section II.4). Before proving this, we need the notion of order statistics.

II.3 Conditional distribution of the arrival times

Definition 2.20 (Order statistic)**.**

Let $Y_1, Y_2, \ldots, Y_n$ be n random variables. We say that $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ are the order statistics corresponding to $Y_1, \ldots, Y_n$ if $Y_{(k)}$ is kth smallest value among the Y_i's. In other words

$$Y(1) \leq Y_{(2)} \leq \ldots \leq Y_{(n)}.$$

As an example, if $n = 3$, $Y_1 = 4$, $Y_2 = 5$ and $Y_3 = 1$, then $Y_{(1)} = 1$, $Y_{(2)} = 4$ and $Y_{(3)} = 5$.

Note that even if the Y_i are independent and identically distributed, the order statistics are NOT independent, nor identically distributed.

Lemma 2.21.

If $Y_1, \dots, Y_n$ are i.i.d. continuous random variables with density f, then the joint density of the order statistics $Y_{(1)}, Y_{(2)}, \dots, Y_{(n)}$ is given by

$$(u_1, u_2, \dots, u_n) \mapsto n! \left(\prod_{i=1}^{n} f(u_i) \right) \mathbb{1}_{\{u_1 < u_2 < \dots < u_n\}}.$$

Proof. Denote $\mathfrak{S}_n$ the set of all permutations of $\{1, 2, \dots, n\}$. Let B be a Borel set in $\mathbb{R}^n$. Then

$$\mathbb{P}\big(\big(Y_{(1)}, \dots, Y_{(n)}\big) \in B\big) = \sum_{\sigma \in \mathfrak{S}_n} \mathbb{P}\big(\big\{\big(Y_{\sigma(1)}, \dots, Y_{\sigma(n)}\big) \in B\big\} \cap \big\{Y_{\sigma(1)} < \dots < Y_{\sigma(n)}\big\}\big).$$

Yet, if $\sigma \in \mathfrak{S}_n$ is deterministic, then $Y_{\sigma(1)}, \dots, Y_{\sigma(n)}$ are alors i.i.d. with density f. Hence

$$\mathbb{P}\big(\big(Y_{(1)}, \dots, Y_{(n)}\big) \in B\big) = \sum_{\sigma \in \mathfrak{S}_n} \int_{\mathbb{R}^n} \mathbb{1}_{\{(u_1, \dots, u_n) \in B\}} \mathbb{1}_{\{u_1 < \dots < u_n\}} f(u_1) \dots f(u_n) du_1 \dots du_n.$$

Finally, since the cardinality of $\mathfrak{S}_n$ equals $n!$,

$$\mathbb{P}\big(\big(Y_{(1)}, \dots, Y_{(n)}\big) \in B\big) = \int_{\mathbb{R}^n} \mathbb{1}_{\{(u_1, \dots, u_n) \in B\}} \underbrace{\left[n! \left(\prod_{i=1}^{n} f(u_i) \right) \mathbb{1}_{\{u_1 < \dots < u_n\}} \right]}_{(*)} du_1 \dots du_n.$$

This being true for any B, we deduce that $(*)$ is the joint density of the order statistics $\big(Y_{(1)}, \dots, Y_{(n)}\big)$. □

In the particular case of uniform random variables, we obtain the following property.

Property 2.22.

i) If $U_1, \dots, U_n$ are i.i.d. r.v. with uniform distribution on $[0, t]$, then the joint density of the order statistics $U_{(1)}, U_{(2)}, \dots, U_{(n)}$ is given by

$$(u_1, u_2, \dots, u_n) \mapsto \frac{n!}{t^n} \mathbb{1}_{\{0 < u_1 < u_2 < \dots < u_n < t\}}.$$

ii) We deduce the following integral

$$\int_{\mathbb{R}^n} \mathbb{1}_{\{0 < u_1 < u_2 < \dots < u_n < t\}} = \frac{t^n}{n!}.$$

Proof. Trivial from Lemma 2.21. □

The following result is very useful to prove the reverse of Proposition 2.18 and thus provides a way of sampling a homogeneous Poisson process. It can also be used to test for homogeneity of a Poisson process.

Proposition 2.23 (Conditional distribution).

Let N be a homogeneous Poisson process with rate $\lambda > 0$ and fix $t > 0$. Let $n \in \mathbb{N}^*$. Given that $N_t = n$, the n first arrival times $(T_1, \dots, T_n)$ have the same distribution as the order statistic corresponding to n independent random variables uniformly distributed on the interval $[0, t]$, that is

$$(T_1, T_2, \dots, T_n) | \{N_t = n\} \overset{(d)}{=} (U_{(1)}, U_{(2)}, \dots, U_{(n)}) \quad \text{where } U_1, \dots, U_n \text{ i.i.d } \sim \mathcal{U}([0, t]),$$

where $\overset{(d)}{=}$ means that both r.v. have the same distribution.

Proof. In exercise. Let $n \in \mathbb{N}^*$ and B be a Borel set.

a) Write the event $\{N_t = n\}$ in terms of the arrival times $(T_i)_i$.

Solution.

The event $\{N_t = n\}$ means that there are exactly n events by time t. This is equivalent to say that the nth event occurred by time t, and the $(n+1)$th after time t. Hence

$$\{N_t = n\} = \{T_n \leq t < T_{n+1}\}.$$

b) Using the definition of a Poisson process, and Lemma 2.19, deduce that

$$\begin{aligned}&\mathbb{P}((T_1, \dots, T_n) \in B | N_t = n) \\ &\quad = \frac{n!}{(\lambda t)^n e^{-\lambda t}} \int_{\mathbb{R}^{n+1}} \mathbb{1}_{\{(t_1, \dots, t_n) \in B\}} \lambda^{n+1} e^{-\lambda t_{n+1}} \mathbb{1}_{\{0 < t_1 < \dots < t_n \leq t < t_{n+1}\}} dt_1 \dots dt_n dt_{n+1}.\end{aligned}$$

Solution.

By definition of the conditional probability,

$$\mathbb{P}((T_1, \dots, T_n) \in B | N_t = n) = \frac{\{\mathbb{P}((T_1, \dots, T_n) \in B) \cap \{N_t = n\}\}}{\mathbb{P}(N_t = n)}.$$

Yet, $N_t \sim \mathcal{P}(\lambda t)$, hence

$$\mathbb{P}(N_t = n) = \frac{(\lambda t)^n}{n!} e^{-\lambda t}.$$

Moreover, $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$. Thus,

$$\mathbb{P}((T_1, \dots, T_n) \in B | N_t = n) = \frac{n!}{(\lambda t)^n e^{-\lambda t}} \mathbb{P}(\{(T_1, \dots, T_n) \in B\} \cap \{T_n \leq t < T_{n+1}\}).$$

Moreover, Lemma 2.19 provides the density of $(T_1, \dots, T_n, T_{n+1})$, directly leading to the desired formula.

c) Deduce that

$$\mathbb{P}((T_1, \dots, T_n) \in B | N_t = n) = \int_{\mathbb{R}^n} \mathbb{1}_{\{(t_1, \dots, t_n) \in B\}} \frac{n!}{t^n} \mathbb{1}_{\{0 < t_1 < \dots < t_n < t\}} dt_1 \dots dt_n.$$

Solution.

By simplification of the λ powers, and by the Fubini-Tonelli theorem, we can integrate first w.r.t t_{n+1}, and obtain

$$\begin{aligned}
&\mathbb{P}\big((T_1,\ldots,T_n)\in B|N_t=n\big)\\
&\quad=\frac{n!}{t^n e^{-\lambda t}}\int_{\mathbb{R}^n}\mathbb{1}_{\{(t_1,\ldots,t_n)\in B\}}\underbrace{\left[\int_t^{+\infty}\lambda e^{-\lambda t_{n+1}}dt_{n+1}\right]}_{e^{-\lambda t}}\mathbb{1}_{\{0<t_1<\ldots<t_n\le t\}}dt_1\ldots dt_n\\
&\quad=\int_{\mathbb{R}^n}\mathbb{1}_{\{(t_1,\ldots,t_n)\in B\}}\frac{n!}{t^n}\mathbb{1}_{\{0<t_1<\ldots<t_n<t\}}dt_1\ldots dt_n.
\end{aligned}$$

d) Conclude.

Solution.

We recognize the density of the order statistic of n i.i.d. uniformly distributed r.v. on $[0,t]$ given in Property 2.22.

□

II.4 Construction of a homogeneous Poisson process

Let us now prove the reverse of Proposition 2.18, which provides an alternative definition of a homogeneous Poisson process.

Theorem 2.24.

Let $(W_i)_{i\ge1}$ be a sequence of i.i.d. random variables with distribution $\mathcal{E}(\lambda)$. Then the point process with arrival times defined for all $n\ge1$ by $T_n=\sum_{i=1}^n W_i$, is a homogeneous Poisson process with rate λ.

This result relies on the following Lemmas.

Lemma 2.25.

Under the assumptions of Theorem 2.24, for all t in $\mathbb{R}_+$ and all n in $\mathbb{N}$,

$$\mathbb{P}(N_t=n)=\frac{(\lambda t)^n}{n!}e^{-\lambda t}.$$

In particular, for all $t>0$, $N_t\sim\mathcal{P}(\lambda t)$.

Proof. Recall that $\{N_t=n\}=\{T_n\le t<T_{n+1}\}$, where

$$\begin{cases} T_n=\displaystyle\sum_{i=1}^n W_i\sim\Gamma(n,\lambda),\\ T_{n+1}=T_n+W_{n+1}\quad\text{with}\quad W_{n+1}\perp\!\!\!\perp T_n,\ \text{and } W_{n+1}\sim\mathcal{E}(\lambda).\end{cases}$$

1. Justify that

$$\mathbb{P}(N_t=n)=\int_{\mathbb{R}^2}\mathbb{1}_{\{s\le t\}}\mathbb{1}_{\{s+w>t\}}\left[\frac{(\lambda s)^{n-1}}{(n-1)!}\lambda e^{-\lambda s}\mathbb{1}_{\{s>0\}}\right]\left(\lambda e^{-\lambda w}\mathbb{1}_{\{w>0\}}\right)dsdw.$$

Solution.

From Definition 1.16, we have

$$\begin{aligned}\mathbb{P}(N_t = n) &= \mathbb{P}(\{T_n \le t\} \cap \{T_n + W_{n+1} > t\}) \\ &= \int_{\mathbb{R}^2} \mathbb{1}_{\{s\le t\}} \mathbb{1}_{\{s+w>t\}} \underbrace{\frac{(\lambda s)^{n-1}}{(n-1)!}\lambda e^{-\lambda s}\mathbb{1}_{\{s>0\}}}_{\text{density of } T_n} \left(\lambda e^{-\lambda w}\mathbb{1}_{\{w>0\}}\right) dsdw.\end{aligned}$$

2. Compute $\mathbb{P}(N_t = n)$.

Solution.

Hence by the Fubini-Tonelli theorem,

$$\begin{aligned}\mathbb{P}(N_t = n) &= \int_0^t \frac{\lambda^n s^{n-1}}{(n-1)!} e^{-\lambda s} \underbrace{\left[\int_{t-s}^{+\infty} \lambda e^{-\lambda w} dw\right]}_{e^{-\lambda(t-s)}} ds \\ &= \frac{\lambda^n}{(n-1)!} e^{-\lambda t} \underbrace{\int_0^t s^{n-1} ds}_{t^n/n} \\ &= \frac{(\lambda t)^n}{n!} e^{-\lambda t}.\end{aligned}$$

□

In order to prove the stationarity and the independence of the increments, we also need the joint distribution of the arrival times.

Lemma 2.26. *Under the assumptions of Theorem 2.24, the arrival times $(T_1, \ldots, T_n)$ have the following density w.r.t. the Lebesgue measure:*

$$(t_1, \ldots, t_n) \mapsto \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}}.$$

Note that we recover the same density as in Lemma 2.19 under the homogeneous Poisson process assumption (phew!).

Proof. "Same" change of variables as in the proof of Lemma 2.19. □

We may now prove Theorem 2.24.

Proof of Theorem 2.24. Notice that Lemmas 2.25 and 2.26 jointly imply that the conditional distribution of $(T_1, \ldots, T_n)|N_t = n$ is the same as the joint distribution of the order statistic $U_{(1)} < \ldots < U_{(n)}$ of i.i.d. r.v. $(U_1, \ldots, U_n)$ with distribution $\mathcal{U}([0,t])$ (exactly same proof as Proposition 2.23).

Let $k \in \mathbb{N}^*$ and $0 = t_0 < t_1 < \ldots < t_k = t$ be a subdivision of $[0,t]$.
Conditionally on $\{N_t = n\}$, since the $(T_1, \ldots, T_n) \overset{(d)}{=} (U_{(1)}, \ldots, U_{(n)})$,

$$N_{t_j} - N_{t_{j-1}} = \sum_{i=1}^n \mathbb{1}_{\{t_{j-1}<T_i\le t_j\}} \overset{(d)}{=} \sum_{i=1}^n \mathbb{1}_{\{t_{j-1}<U_{(i)}\le t_j\}} = \sum_{i=1}^n \mathbb{1}_{\{t_{j-1}<U_i\le t_j\}}.$$

Then $(N_{t_1}, N_{t_2} - N_{t_1}, \dots, N_{t_k} - N_{t_{k-1}})$ has a multinomial distribution with parameters $(n, (p_j)_{1 \le j \le k})$ where

$$p_j = \mathbb{P}(t_{j-1} < U_1 \le t_j) = \frac{t_j - t_{j-1}}{t}, \qquad \text{since } U_1 \sim \mathcal{U}([0,t[).$$

Hence, for all integers $n_1, \dots, n_k$ such that $\sum_{j=1}^k n_j = n$,

$$\begin{aligned}\mathbb{P}\big(N_{t_1} = n_1, N_{t_2} - N_{t_1} = n_2, \dots, N_{t_k} - N_{t_{k-1}} = n_k | N_t = n\big) &= \frac{n!}{n_1! \dots n_k!} p_1^{n_1} \dots p_k^{n_k} \\ &= \frac{n!}{t^n} \prod_{j=1}^k \frac{(t_j - t_{j-1})^{n_j}}{n_j!}.\end{aligned}$$

We thus obtain the unconditional probability

$$\begin{aligned}&\mathbb{P}\big(N_{t_1} = n_1, N_{t_2} - N_{t_1} = n_2, \dots, N_{t_k} - N_{t_{k-1}} = n_k\big) \\ &\quad = \mathbb{P}\big(N_{t_1} = n_1, N_{t_2} - N_{t_1} = n_2, \dots, N_{t_k} - N_{t_{k-1}} = n_k | N_t = n\big)\, \mathbb{P}(N_t = n) \\ &\quad = \frac{n!}{t^n} \prod_{j=1}^k \left[\frac{(t_j - t_{j-1})^{n_j}}{n_j!}\right] \times \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\ &\quad = \prod_{j=1}^k \left(\frac{[\lambda(t_j - t_{j-1})]^{n_j}}{n_j!} e^{-\lambda(t_j - t_{j-1})}\right),\end{aligned}$$

since

$$\lambda^n = \lambda^{\sum_{j=1}^k n_j} = \prod_{j=1}^k \lambda^{n_j} \qquad \text{and} \qquad e^{-\lambda t} = e^{-\lambda \sum_{j=1}^k (t_j - t_{j-1})} = \prod_{j=1}^k e^{-\lambda(t_j - t_{j-1})}.$$

We recognize the joint distribution of k independent r.v. with respective distribution $\mathcal{P}(\lambda(t_j - t_{j-1}))$. This directly implies points (2) (independence of the increments) and (3) (the number of points in an interval $(t_{j-1}, t_j]$ has a Poisson distribution with parameter λ times the length $(t_j - t_{j-1})$ of the interval) in Definition 2.11. □

A possible generalization of Poisson processes is to consider a point process for which the times between successive events are independent and identically distributed (but not necessarily exponentially distributed). Such a point process is called *renewal process*.

III Divisibility of homogeneous Poisson processes

III.1 Two states divisibility

Proposition 2.27.

Consider a homogeneous Poisson process N with rate $\lambda > 0$ and assume that each time an event occurs it is classified as either a type I with probability p or a type II event with probability $(1-p)$ independently of all other events. Let N^{I} and N^{II} respectively denote the point processes corresponding to type I and type II events.

Then,

1. *For all $t > 0$, the total number of events occurring in $[0, t)$ equals $N_t = N_t^{\mathrm{I}} + N_t^{\mathrm{II}}$.*
2. *N^{I} and N^{II} are both Poisson processes having respective rates $p\lambda$ and $(1-p)\lambda$.*
3. *Furthermore, the two processes N^{I} and N^{II} are independent.*

This result lies on the following key point. The classification of each point being independent on when it occurs, we can say that the conditional distribution of N_t^{I} given that $N_t = n$ has a binomial distribution with parameters (n, p), that is

$$N_t^{\mathrm{I}} | \{N_t = n\} \sim \mathcal{B}(n, p). \tag{2.5}$$

Proof. Point 1 is trivial. Let us prove 2 and 3.

- The independence of the increments of N^{I} and N^{II} comes from the fact that N is a Poisson process. Hence the number of points in disjoint intervals are independent. Moreover, the type is decided independently on the point process.
- The stationarity of the increments of N^{I} and N^{II} also also comes from the fact that N is a Poisson process. Indeed, the number of points of N in two intervals with same length has the same distribution. Moreover, once again, the classification does not depend on the times the event occurs.
- **Poisson distribution and $\boldsymbol{N^{\mathrm{I}} \perp\!\!\!\perp N^{\mathrm{II}}}$:** Let $t > 0$ and $k, l \in \mathbb{N}$. Then,

$$\begin{aligned}\mathbb{P}\left(N_t^{\mathrm{I}} = k; N_t^{\mathrm{II}} = l\right) &= \mathbb{P}\left(N_t^{\mathrm{I}} = k; N_t = k + l\right) \\ &= \mathbb{P}\left(N_t^{\mathrm{I}} = k \middle| N_t = k + l\right) \mathbb{P}\left(N_t = k + l\right).\end{aligned}$$

Yet, from the key point stated above, the conditional distribution of N_t^{I} given $\{N_t = k + l\}$ is a binomial $\mathcal{B}(k + l, p)$. Moreover, $N_t \sim \mathcal{P}(\lambda t)$. Hence

$$\begin{aligned}\mathbb{P}\left(N_t^{\mathrm{I}} = k; N_t^{\mathrm{II}} = l\right) &= \frac{(k+l)!}{k!\ l!}\, p^k (1-p)^l\, \frac{(\lambda t)^{k+l}}{(k+l)!} e^{-\lambda t} \\ &= \underbrace{\left(\frac{(p\lambda t)^k}{k!} e^{-p\lambda t}\right)}_{\mathbb{P}\left(N_t^{\mathrm{I}} = k\right)} \underbrace{\left(\frac{((1-p)\lambda t)^l}{l!} e^{-(1-p)\lambda t}\right)}_{\mathbb{P}\left(N_t^{\mathrm{II}} = l\right)}\end{aligned}$$

We deduce that

$$\begin{cases} N_t^{\mathrm{I}} \sim \mathcal{P}(p\lambda t) \\ N_t^{\mathrm{II}} \sim \mathcal{P}((1-p)\lambda t) \\ N_t^{\mathrm{I}} \perp\!\!\!\perp N_t^{\mathrm{II}} \end{cases}$$

□

Example 2.28.

Clients of an insurance company call at a Poisson process rate of ten per week. In 10% of the cases, it is to sign up a new policy, and in the remaining 90%, to report a claim. What is the probability that there will be no new customers in February?

Solution.

The number of new customers during February follows a Poisson distribution with parameter 4 (weeks) $\times$ 10 (λ) $\times$ 1/10 (p) $= 4$ so the answer is e^{-4}.

The reverse of Proposition 2.27 is true.

Proposition 2.29.

If N^{I} and N^{II} are two independent Poisson processes with respective rates λ_{I} and λ_{II}, then the counting process $N = (N_t)_{t \geq 0}$ defined for all t in $\mathbb{R}_+$ by

$$N_t = N_t^{\mathrm{I}} + N_t^{\mathrm{II}}$$

is also a Poisson process with rate $\lambda_{\mathrm{I}} + \lambda_{\mathrm{II}}$.

Proof. Facultative homework. □

III.2 Infinite divisibility

The following generalization of Proposition 2.27 can be easily proved by mathematical induction.

Proposition 2.30.

Let N be a homogeneous Poisson process with intensity λ. Assume each time an event occurs is associated to one of M types in $\{1, \ldots, M\}$ with probability $1/M$ independently of all other events. For all m in $\{1, \ldots, M\}$, denote N^m the point process associated to events of type m. Then $N^1, \ldots, N^m$ are m i.i.d. homogeneous Poisson processes with intensity λ/M.

Definition 2.31.

A random variable X is said to be infinitely divisible if for all M in N^*, there exists M i.i.d. random variables $Y_1, \ldots, Y_M$ such that X has the same distribution as $Y_1 + \ldots + Y_M$.

Property 2.32.

Random variables with distribution $\mathcal{P}(\lambda)$, $\mathcal{N}(\mu, \sigma^2)$, $\mathcal{E}(\lambda)$ or $\Gamma(n, \lambda)$ are infinitely divisible.

Property 2.33.

A homogeneous Poisson processes is infinitely divisible.

Chapter 3

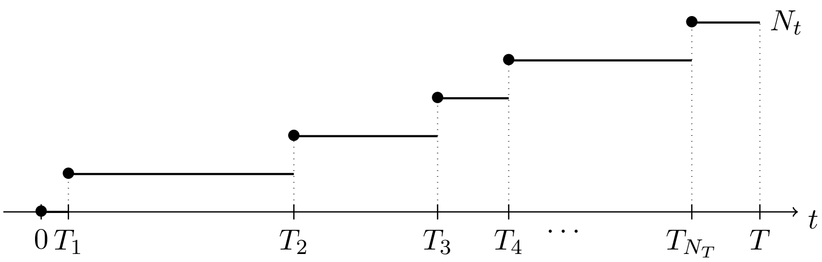
Statistics for homogeneous Poisson processes

In this part, we do some statistical inference for homogeneous Poisson processes. We observe a random point process $N = (N_t)_{t\in\mathbb{R}_+}$ with arrival times $(T_n)_{n\in\mathbb{N}^*}$. We assume that N is a homogeneous Poisson process with *unknown* intensity $\lambda > 0$. We aim at estimating λ. We also aim at constructing (asymptotic) confidence intervals for λ, and statistical tests.

I Introduction

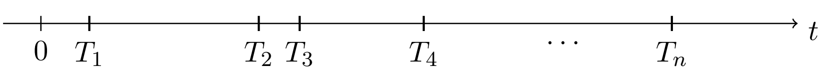
There are two types of observation for a point process:

- Either the observation interval $[0, T]$ is fixed, and we observe the process on a fixed interval, that is we observe $(N_t)_{t\in[0,T]}$.



  In this case, the number of events occurring in this interval N_T is random.

- Either n is fixed, and we observe the process up to the nth event, that is we observe $(T_1, \dots, T_n)$.



  In this case, the observation interval/window $[0, T_n]$ is random.

II Fixed window, random number of points

In this section T is fixed, and we observe a homogeneous Poisson process N with (unknown) intensity $\lambda > 0$ on the interval $[0, T]$. In particular, the number of "events" that have occurred by time T, that is N_T, is random.

II.1 Maximum Likelihood Estimator

Heuristic: Recall that, if X is a random variable with density f_θ and c.d.f. F_θ. Then, its likelihood in θ equals

$$\mathcal{L}(X,\theta) = f_\theta(X),$$

where,

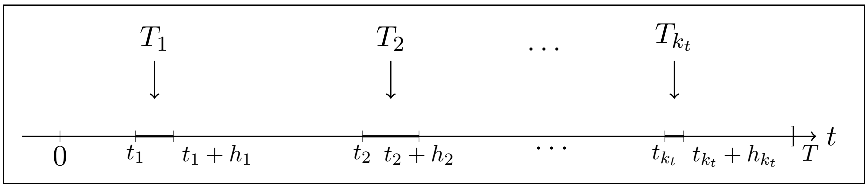
$$f_\theta(x) = F'_\theta(x) = \lim_{h\to 0} \frac{F_\theta(x+h) - F_\theta(x)}{h} = \lim_{h\to 0} \frac{\mathbb{P}_\theta(X \in (x; x+h])}{h}.$$

In the point process case: Recall that in the "fixed-T" case, $N_T = \sum_{i=1}^{+\infty} \mathbb{1}_{\{T_i \leq T\}}$ is random. Similarly to the proof of Lemma 2.19, consider a strictly increasing sequence $(t_i)_{i\in\mathbb{N}^*}$ in $\mathbb{R}_+$, and denote

$$k_t = \sum_{i=1}^{+\infty} \mathbb{1}_{\{t_i \leq T\}} \qquad \text{s.t.} \qquad 0 =: t_0 < t_1 < \ldots < t_{k_t} < T < t_{k_t+1} < \ldots.$$

Hence, k_t counts the number of t_i's that belong to $[0,T]$.
Introduce for $h = (h_1, \ldots, h_n)$ small enough such that all $(t_i, t_i + h_i]$ are disjoint (and $t_k + h_k \leq T$),

$$B_h = \{1 \text{ point in each } (t_i, t_i + h_i] \text{ and } 0 \text{ points elsewhere in } [0,T]\}.$$



The main difference with A_h in Lemma 2.19 is that now n in not fixed anymore, and we need to consider k_t which is the number of (t_i) which belong to the observation interval $[0,T]$.

Heuristically, as in the proof of Lemma 2.19,

$$f((t_i)_{i\in\mathbb{N}^*} \cap [0,T]) = \lim_{\forall i,\ h_i \to 0} \frac{\mathbb{P}(B_h)}{h_1 \ldots h_{k_t}}.$$

Since the intervals appearing in B_h are disjoints, by independence of the increments,

$$\begin{aligned}
\mathbb{P}(B_h) &= \left[\prod_{i=1}^{k_t} \mathbb{P}(1 \text{ point in } (t_i, t_i + h_i])\right] \times \mathbb{P}(0 \text{ points elsewhere}) \\
&\overset{(*)}{=} \left[\prod_{i=1}^{k_t} \lambda h_i e^{-\lambda h_i}\right] \times e^{-\lambda(T - \sum_{i=1}^{k_t} h_i)} \\
&= \lambda^{k_t} h_1 \ldots h_{k_t} e^{-\lambda T}.
\end{aligned}$$

with $^{(*)}$ is because "elsewhere" is of length $T - \sum_{i=1}^{k_t} h_i$.

Finally,

$$f((t_i)_{i\in\mathbb{N}^*} \cap [0,T]) = \lim_{\forall i,\ h_i \to 0} \lambda^{k_t} e^{-\lambda T} = \lambda^{k_t} e^{-\lambda T}.$$

Proposition 3.1.

Let N be a homogeneous Poisson process with (unknown) intensity λ. Consider the number N_T of "events" that have occurred in $[0,T]$, dans denote $(T_1,\ldots,T_{N_T})$ the corresponding arrival times. Then likelihood of the trajectory of N is equal to

$$\mathcal{L}\left((N_t)_{t\in[0,T]};\lambda\right) = \lambda^{N_T} e^{-\lambda T}.$$

Hence, the Maximum Likelihood Estimator (MLE) of λ is

$$\hat{\lambda}_T = \frac{N_T}{T}.$$

Notice that the likelihood can be expressed as follows:

$$\mathcal{L}\left((N_t)_{t\in[0,T]};\lambda\right) = \left[N_T!\left(\prod_{i=1}^{N_T}\frac{1}{T}\mathbb{1}_{\{T_i\in[0,T]\}}\right)\mathbb{1}_{\{T_1<\ldots<T_{N_T}\}}\right]\times\left[\frac{(\lambda T)^{N_T}e^{-\lambda T}}{N_T!}\right].$$

We recognize the product of

- the (conditional) likelihood of $(T_1,\ldots,T_{N_T})|N_T$, that is, according to Proposition 2.23, the likelihood of the order statistic corresponding to N_T independent random variables uniformly distributed on the interval $[0,T]$ given N_T,
- and the likelihood of N_T which has a Poisson distribution $\mathcal{P}(\lambda T)$.

Proof. In exercise. Compute the MLE of λ.
The expression of the likelihood is admitted here (c.f. heuristic above).

Solution.

We deduce that the log-likelihood equals

$$\ell\left((N_t)_{t\in[0;T]};\lambda\right) = N_T\ln(\lambda) - \lambda T.$$

Hence

$$\frac{\partial \ell}{\partial\lambda}\left((N_t)_{t\in[0;T]};\lambda\right) = \frac{N_T}{\lambda} - T = 0 \qquad\Leftrightarrow\qquad \lambda = \frac{N_T}{T}.$$

Moreover, N_T/T is a maximum since the log-likelihood function is concave. Indeed,

$$\frac{\partial^2 \ell}{\partial\lambda^2}\left((N_t)_{t\in[0;T]};\lambda\right) = \frac{-N_T}{\lambda^2} \leq 0 \quad \text{a.s.}$$

□

II.2 Non-asymptotic properties of the MLE

Proposition 3.2.

Let $(N_t)_{t\in[0,T]}$ be a Poisson process with rate $\lambda>0$ observed on $[0,T]$ and consider the MLE $\hat{\lambda}_T = N_T/T$. Since $N_T\sim\mathcal{P}(\lambda T)$,

$$\mathbb{E}\left[\hat{\lambda}_T\right] = \lambda \qquad \textit{and} \qquad \mathrm{Var}\left(\hat{\lambda}_T\right) = \frac{\lambda}{T}.$$

We deduce that $\hat{\lambda}_T$ is an unbiased and efficient estimator of λ.

Proof. In exercise.

Solution.

Since $N_T \sim \mathcal{P}(\lambda T)$, $\mathbb{E}[N_T] = \lambda T$ and $\mathrm{Var}(N_T) = \lambda T$. Hence, on the one hand, by linearity of the expectation,

$$\mathbb{E}\left[\hat{\lambda}_T\right] = \frac{\mathbb{E}[N_T]}{T} = \lambda.$$

We deduce that $\hat{\lambda}_T$ is an unbiased estimator of λ. On the other hand, by property of the variance,

$$\mathrm{Var}\left(\hat{\lambda}_T\right) = \frac{\mathbb{E}[N_T]}{T^2} = \frac{\lambda}{T}.$$

Moreover, one can compute the Fisher information:

$$I_T(\lambda) = \mathbb{E}_\lambda\left[-\frac{\partial^2}{\partial\lambda^2}\ell\left((N_t)_{t\in[0,T]};\lambda\right)\right] = \mathbb{E}_\lambda\left[\frac{N_T}{\lambda^2}\right] = \frac{T}{\lambda}.$$

Thus the Cramér-Rao bound equals

$$\mathcal{B}_T(\lambda) = \frac{1}{I_T(\lambda)} = \frac{\lambda}{T} = \mathrm{Var}\left(\hat{\lambda}_T\right).$$

We deduce that $\hat{\lambda}_T$ is efficient.

□

Moreover, to construct statistical tests with uniform asymptotic level, we need the following lemma.

Lemma 3.3.

Let N be a homogeneous Poisson process with (unknown) intensity $\lambda > 0$ observed on the interval $[0,T]$ and consider the MLE $\hat{\lambda}_T = N_T/T$ of λ.
Then, for all (fixed) $x \geq 0$,

$$\lambda \mapsto \mathbb{P}_\lambda\left(\hat{\lambda}_T \geq x\right)$$

is a non-decreasing function on $\mathbb{R}_+^*$.

Proof. In exercise. Fix $x \geq 0$ and denote

$$g : \lambda > 0 \mapsto \mathbb{P}_\lambda\left(\hat{\lambda}_T \geq x\right).$$

1. Assume $x = 0$. Compute $g(\lambda)$ for all $\lambda > 0$ and conclude in this case.

   Solution.

   If $x = 0$, g is constant equal to 1, since for all $\lambda > 0$, $\mathbb{P}\left(\hat{\lambda}_T \geq 0\right) = \mathbb{P}(N_T \geq 0) = 1$.

2. Now assume $x > 0$. Justify that for all $\lambda > 0$,

$$g(\lambda) = \sum_{k=\lceil xT\rceil}^{+\infty} \frac{(\lambda T)^k}{k!}e^{-\lambda T},$$

   where $\lceil\cdot\rceil$ denotes the ceiling function.

   Solution.

   Note that

$$g(\lambda) = \mathbb{P}\left(\hat{\lambda}_T \geq x\right) = \mathbb{P}(N_T \geq xT) = \mathbb{P}(N_T \geq \lceil xT\rceil),$$

   since N_T is an integer. The expression of $g(\lambda)$ is then immediate since $N_T \sim \mathcal{P}(\lambda T)$.

3. Deduce that
$$g'(\lambda) = \frac{(\lambda T)^{\lceil xT \rceil - 1}}{(\lceil xT \rceil - 1)!} T e^{-\lambda T}.$$
Hint: Use the method of differences, that is telescoping series.

Solution.

We obtain
$$\begin{aligned} g'(\lambda) &= \left[\sum_{k=\lceil xT \rceil}^{+\infty} \frac{\lambda^{k-1}T^k}{(k-1)!} e^{-\lambda T}\right] - \left[\sum_{k=\lceil xT \rceil}^{+\infty} \frac{(\lambda T)^k}{k!} T e^{-\lambda T}\right] \\ &\overset{(*)}{=} \left(\sum_{k=\lceil xT \rceil - 1}^{+\infty} \frac{(\lambda T)^k}{k!} - \sum_{k=\lceil xT \rceil}^{+\infty} \frac{(\lambda T)^k}{k!}\right) T e^{-\lambda T} \\ &\overset{(\dagger)}{=} \frac{(\lambda T)^{\lceil xT \rceil - 1}}{(\lceil xT \rceil - 1)!} T e^{-\lambda T}, \end{aligned}$$
with $^{(*)}$ by change of variable $k \leftarrow k-1$ (since $\lceil xT \rceil \geq 1$), and $^{(\dagger)}$ by the difference method.

4. Conclude.

Solution.

We deduce that $g'(\lambda) > 0$ for all $\lambda > 0$, that is g is (strictly) increasing when $x > 0$.

□

II.3 Asymptotic distribution of the MLE

Theorem 3.4.

Let $(N_t)_{t \in [0,T]}$ be a Poisson process with rate $\lambda > 0$ observed on $[0, T]$ and consider the MLE $\hat{\lambda}_T = N_T / T$.

1. *LLN-type result: $\hat{\lambda}_T$ is a consistent estimator of λ, that is*
$$\frac{N_T}{T} \xrightarrow[T \to +\infty]{\mathbb{P}} \lambda.$$

2. *CLT-type result: Moreover, it is asymptotically gaussian, that is*
$$\sqrt{T}\left(\frac{N_T}{T} - \lambda\right) \xrightarrow[T \to +\infty]{\mathcal{L}} \mathcal{N}(0, \lambda).$$

Note that it is also possible to prove that $N_T \xrightarrow[T \to +\infty]{a.s.} +\infty$ and deduce the "strong" version of the LLN-type result:
$$\frac{N_T}{T} \xrightarrow[T \to +\infty]{a.s.} \lambda.$$
This result is based on the following inequality: $\dfrac{T_{N_T}}{N_T} \leq \dfrac{T}{N_T} < \dfrac{T_{N_T+1}}{N_T}$.

Proof. In exercise.

1. Note that we cannot apply the classical Law of Large Numbers since the index in the limit is not a countable sequence. Apply Chebychev's inequality to prove the consistency of $\hat{\lambda}_T$.

   Solution.

   Since $\mathbb{E}\left[\hat{\lambda}_T\right] = \lambda$ and $\mathrm{Var}\left(\hat{\lambda}_T\right) = \lambda/T < +\infty$, for all $\varepsilon > 0$,

$$\mathbb{P}\left(\left|\hat{\lambda}_T - \lambda\right| \geq \varepsilon\right) \leq \frac{\mathrm{Var}\left(\hat{\lambda}_T\right)}{\varepsilon^2} = \frac{\lambda}{T\varepsilon^2} \underset{T\to+\infty}{\longrightarrow} 0.$$

2. Prove that the characteristic function of $Z_T := \sqrt{T}(N_T/T - \lambda)$ equals for all $u \in \mathbb{R}$,

$$\mathbb{E}\left[e^{iuZ_T}\right] = \exp\left(\lambda T\left[e^{\frac{iu}{\sqrt{T}}} - 1 - \frac{iu}{\sqrt{T}}\right]\right).$$

   Hint: Use Proposition 1.2.

   Solution.

   Let $u \in \mathbb{R}$.

$$\mathbb{E}\left[e^{iuZ_T}\right] = \mathbb{E}\left[e^{iu\sqrt{T}\left(\frac{N_T}{T} - \lambda\right)}\right] = \mathbb{E}\left[e^{i\frac{u}{\sqrt{T}}N_T}\right] e^{-iu\lambda\sqrt{T}}$$

   Yet $N_T \sim \mathcal{P}(\lambda T)$. Hence, by Proposition 1.2, its characteristic function computed in $u/\sqrt{T}$ equals

$$\mathbb{E}\left[e^{i\frac{u}{\sqrt{T}}N_T}\right] = \exp\left(\lambda T\left[e^{i\frac{u}{\sqrt{T}}} - 1\right]\right).$$

   Hence, we directly obtain

$$\mathbb{E}\left[e^{iuZ_T}\right] = \exp\left(\lambda T\left[e^{\frac{iu}{\sqrt{T}}} - 1 - \frac{iu}{\sqrt{T}}\right]\right).$$

3. Prove that

$$\lim_{T\to+\infty} \mathbb{E}\left[e^{iuZ_T}\right] = e^{-\frac{\lambda u^2}{2}}.$$

   Hint: Use Taylor's expansion.

   Solution.

   By Taylor's expansion, as $x \to 0$,

$$\exp(x) = 1 + x + \frac{x^2}{2} + o\left(x^2\right).$$

   Hence as $T \to +\infty$,

$$\mathbb{E}\left[e^{iuZ_T}\right] = \exp\left(\lambda T\left[\frac{-u^2}{2T} + o\left(\frac{1}{T}\right)\right]\right) = \exp\left(-\frac{\lambda u^2}{2} + o(1)\right) \underset{T\to+\infty}{\longrightarrow} \exp\left(-\frac{\lambda u^2}{2}\right).$$

4. Deduce the asymptotic distribution of Z_T. *Hint:* Recall that the characteristic function of a gaussian r.v. Z with distribution $\mathcal{N}(m, \sigma^2)$ equals for all $u \in \mathbb{R}$, $\mathbb{E}\left[e^{iuZ}\right] = \exp\left(imu - \frac{u^2\sigma^2}{2}\right)$.

Solution.

We recognize the characteristic function of a $\mathcal{N}(0,\lambda)$, which ends the proof of Theorem 3.4.

□

II.4 Confidence intervals and statistical tests

Fix $T > 0$ and let α in $(0,1)$. We observe a homogeneous Poisson process with unknown intensity $\lambda > 0$ on $[0,T]$ and we aim at constructing confidence intervals for λ with confidence level $1-\alpha$, or at testing, for instance,

$$\mathcal{H}_0 : \lambda \leq \lambda_0 \quad \text{against} \quad \mathcal{H}_1 : \lambda > \lambda_0, \qquad \text{at level } \alpha.$$

From Theorem 3.4 and Slutsky's Lemma, for all α in $(0,1)$, we construct an asymptotic confidence interval for λ with confidence level $1-\alpha$, that is

$$\left[\frac{N_T}{T} - \frac{\sqrt{N_T}}{T} z_{1-\alpha/2}; \frac{N_T}{T} + \frac{\sqrt{N_T}}{T} z_{1-\alpha/2}\right],$$

where $z_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of a standard gaussian distribution $\mathcal{N}(0,1)$.

Proof. In exercise.

Solution.

By Slutsky's Lemma,

$$\frac{T}{\sqrt{N_T}}\left(\frac{N_T}{T} - \lambda\right) = \underbrace{\sqrt{\frac{\lambda}{N_T/T}}}_{(a)} \times \underbrace{\sqrt{\frac{T}{\lambda}}\left(\frac{N_T}{T} - \lambda\right)}_{(b)} \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}(0,1),$$

Since by the LLN-type result in Theorem 3.4 and by continuity of $x \mapsto \sqrt{\lambda/x}$ on $\mathbb{R}_+^*$,

$$(a) \xrightarrow[T\to+\infty]{\mathbb{P}} 1,$$

and by the CLT-type result in Theorem 3.4,

$$(b) \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

We deduce that

$$\mathbb{P}\left(\frac{N_T}{T} - \frac{\sqrt{N_T}}{T} z_{1-\alpha/2} \leq \lambda \leq \frac{N_T}{T} + \frac{\sqrt{N_T}}{T} z_{1-\alpha/2}\right) = \mathbb{P}\left(-z_{1-\alpha/2} \leq \frac{T}{\sqrt{N_T}}\left(\frac{N_T}{T} - \lambda\right) \leq z_{1-\alpha/2}\right)$$
$$\xrightarrow[T\to+\infty]{} \mathbb{P}\left(-z_{1-\alpha/2} \leq Z \leq z_{1-\alpha/2}\right) = 1-\alpha,$$

where Z denotes a standard gaussian r.v.

□

Moreover, one can construct different test of

$$\mathcal{H}_0 : \lambda \leq \lambda_0 \quad \text{against} \quad \mathcal{H}_1 : \lambda > \lambda_0, \qquad \text{at level } \alpha.$$

Test 1: Reject $\mathcal{H}_0$ if $\hat{\lambda}_T \geq \lambda_0 + \sqrt{\frac{\lambda_0}{T}} z_{1-\alpha}$, with $\begin{cases} \text{test statistic: } Z_T^{(1)} = \sqrt{T}\frac{\hat{\lambda}_T - \lambda_0}{\sqrt{\lambda_0}}, \\ p\text{-value: } 1 - \Phi(Z_T^{(1),obs}), \end{cases}$

Test 2: Reject $\mathcal{H}_0$ if $\hat{\lambda}_T \geq \lambda_0 + \sqrt{\frac{\hat{\lambda}_T}{T}} z_{1-\alpha}$, with $\begin{cases} \text{test statistic: } Z_T^{(2)} = \sqrt{T}\frac{\hat{\lambda}_T - \lambda_0}{\sqrt{\hat{\lambda}_T}}, \\ p\text{-value: } 1 - \Phi(Z_T^{(2),obs}), \end{cases}$

where $z_{1-\alpha}$ and Φ respectively denote the $(1-\alpha)$-quantile and the c.d.f. of a standard gaussian distribution $\mathcal{N}(0,1)$.

One can prove that both Test 1 and Test 2 are (uniformly) asymptotically of level α, that is

$$\limsup_{n\to+\infty} \left(\sup_{\lambda \leq \lambda_0} \{\mathbb{P}_\lambda(\text{reject } \mathcal{H}_0)\} \right) \leq \alpha.$$

This results relies on Lemma 3.3.

Proof. Facultative homework. □

III Fixed number of points, random observation time

In this section n is fixed, and we observe a homogeneous Poisson process N with (unknown) intensity $\lambda > 0$ up to the nth arrival times. In particular, we observe the n first arrival times $(T_1, \ldots, T_n)$ or equivalently the n first interarrival times $(W_1, \ldots, W_n)$. Then, according to Proposition 2.18, the $(W_i)_{i\geq 1}$ are i.i.d. with distribution $\mathcal{E}(\lambda)$ which makes the study easier. In particular, T_n is $\Gamma(n, \lambda)$-distributed.

III.1 Maximum Likelihood Estimator (MLE)

Proposition 3.5.

Let N be a homogeneous Poisson process with (unknown) intensity $\lambda > 0$. Consider the n first arrival times $(T_1, \ldots, T_n)$ of N. Then likelihood of the trajectory of N is equal to

$$\mathcal{L}\left((T_1, \ldots, T_n); \lambda\right) = \lambda^n e^{-\lambda T_n}.$$

Hence, the Maximum Likelihood Estimator (MLE) of λ is

$$\hat{\lambda}_n = \frac{n}{T_n}.$$

Proof. In exercise.

1. Recall the joint density of $(T_1, \ldots, T_n)$, and deduce its likelihood.

   Solution.

   According to Lemma 2.19, we know that the density of $(T_1, \ldots, T_n)$ is

   $$(t_1, \ldots, t_n) \mapsto \lambda^n e^{-\lambda t_n} \mathbb{1}_{\{0<t_1<\ldots<t_n\}}.$$

   Hence its likelihood equals

   $$\mathcal{L}\left((T_1, \ldots, T_n); \lambda\right) = \lambda^n e^{-\lambda T_n} \mathbb{1}_{\{0<T_1<\ldots<T_n\}},$$

   with $\mathbb{1}_{\{0<T_1<\ldots<T_n\}} = 1$ a.s.

2. Compute the MLE of λ.

   Solution.

   We deduce that the log-likelihood equals

   $$\ell\left((T_1, \ldots, T_n); \lambda\right) = n \ln(\lambda) - \lambda T_n.$$

   Hence

   $$\frac{\partial \ell}{\partial \lambda}\left((T_1, \ldots, T_n); \lambda\right) = \frac{n}{\lambda} - T_n = 0 \qquad \Leftrightarrow \qquad \lambda = \frac{n}{T_n}.$$

   Moreover, n/T_n is a maximum since the log-likelihood function is concave. Indeed,

   $$\frac{\partial^2 \ell}{\partial \lambda^2}\left((T_1, \ldots, T_n); \lambda\right) = \frac{-n}{\lambda^2} < 0.$$

$\square$

Note that we could recover this result using that knowing the point process is also equivalent to knowing the interarrival times which are i.i.d. with $\mathcal{E}(\lambda)$ distribution. Indeed,

$$\mathcal{L}((W_1,\dots,W_n);\lambda) = \prod_{i=1}^{n}\left(\lambda e^{-\lambda W_i}\mathbb{1}_{\{W_i>0\}}\right) = \lambda^n e^{-\lambda\sum_{i=1}^n W_i} = \lambda^n e^{-\lambda T_n}.$$

III.2 Non-asymptotic properties of the MLE

Lemma 3.6.

Let N be a homogeneous Poisson process with intensity $\lambda > 0$ and $n \geq 2$ a fixed integer. Then, the MLE of λ satisfies

$$\mathbb{E}\left[\hat{\lambda}_n\right] = \frac{n}{n-1}\lambda.$$

In particular, from Theorem 3.9, we can deduce that in the case "fixed n", the Maximum Likelihood Estimator $\hat{\lambda}_n$ of λ is biased, and asymptotically unbiased.

However, we can consider

$$\tilde{\lambda} = \frac{n-1}{T_n},$$

and prove that $\tilde{\lambda}$ is an unbiased estimator of λ, but not efficient.

Proof. Worksheet 3, exercise 1. □

Moreover, to construct statistical tests with uniform asymptotic level, we need the following lemma.

Lemma 3.7.

Let N be a homogeneous Poisson process with (unknown) intensity $\lambda > 0$ observed up to the nth first arrival time T_n and consider the MLE $\hat{\lambda}_n = n/T_n$ of λ.
Then, for all (fixed) $t > 0$,

$$\lambda \mapsto \mathbb{P}_\lambda\left(\hat{\lambda}_n \geq t\right)$$

is a non-decreasing function.

Proof. In exercise. Consider $\lambda \leq \mu$ and prove that $\mathbb{P}_\lambda\left(\hat{\lambda}_n \geq t\right) \leq \mathbb{P}_\mu\left(\hat{\lambda}_n \geq t\right)$. *Hint:* Use Proposition 1.18.

Solution.

Note that $\hat{\lambda}_n = n/T_n$ with $T_n \sim \Gamma(n,\lambda)$. Hence, by Proposition 1.18,

$$2\lambda T_n \sim \Gamma(n, 1/2) \overset{(d)}{=} \chi^2(2n).$$

Therefore, for $\lambda \leq \mu$,

$$\begin{aligned}
\mathbb{P}_\lambda\left(\hat{\lambda}_n \geq t\right) &= \mathbb{P}_\lambda\left(T_n \leq \frac{n}{t}\right) \\
&= \mathbb{P}_\lambda\left(2\lambda T_n \leq \frac{2\lambda n}{t}\right) \\
&= \mathbb{P}\left(X \leq \frac{2\lambda n}{t}\right) && \text{where } X \sim \chi^2(2n) \\
&\overset{(*)}{\leq} \mathbb{P}\left(X \leq \frac{2\mu n}{t}\right) && \text{since } \lambda \leq \mu \\
&= \mathbb{P}_\mu\left(2\mu T_n \leq \frac{2\mu n}{t}\right) \\
&= \mathbb{P}_\mu\left(\hat{\lambda}_n \geq t\right).
\end{aligned}$$

where $^{(*)}$ holds from the fact that a c.d.f. is (always) a non-decreasing function.

□

III.3 Asymptotic distribution of the MLE

Before studying the asymptotic properties of the MLE, recall the univariate delta method.

Lemma 3.8 (Univariate delta method)**.**

Let $(X_n)_{n\in\mathbb{N}}$ be a sequence of real-valued random variables, Y be a real-valued random variable and θ be in $\mathbb{R}$ such that

$$r_n(X_n - \theta) \xrightarrow[n\to+\infty]{\mathcal{L}} Y,$$

where (r_n) is a sequence of non-negative real numbers that tends toward infinity. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function at θ. Then,

$$f(X_n) \xrightarrow[n\to+\infty]{\mathbb{P}} f(\theta) \qquad \text{and} \qquad r_n\left(f(X_n) - f(\theta)\right) \xrightarrow[n\to+\infty]{\mathcal{L}} f'(\theta)Y.$$

This allows us to deduce the asymptotic behavior of the MLE.

Theorem 3.9.

Let N be a homogeneous Poisson process with (unknown) intensity $\lambda > 0$ observed up to the nth first arrival time T_n and consider the MLE $\hat{\lambda}_n = n/T_n$ of λ.

1. **LLN-type result:** *$\hat{\lambda}_n$ is a consistent estimator of λ, that is*

$$\frac{n}{T_n} \xrightarrow[n\to+\infty]{\mathbb{P}} \lambda.$$

2. **CLT-type result:** *Moreover, it is asymptotically gaussian, that is*

$$\sqrt{n}\left(\frac{n}{T_n} - \lambda\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \lambda^2).$$

Comment 3.10. Note that here, the asymptotic variance equals λ^2.

Proof. In exercise.

1. Apply the CLT to the interarrival times $(W_i)_{i\geq 1}$, and prove that

$$\sqrt{n}\left(\frac{1}{\hat{\lambda}_n}-\frac{1}{\lambda}\right)\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}\left(0,\frac{1}{\lambda^2}\right). \tag{3.1}$$

Solution.

Recall that $\hat{\lambda}_n = n/T_n$ where

$$\frac{T_n}{n}=\frac{1}{n}\sum_{i=1}^{n}W_i=\bar{W}_n.$$

Applying the CLT to the sequence $(W_i)_{i\geq 1}$ which are i.i.d. exponential r.v. such that

$$\mathbb{E}[|W_1|]=\mathbb{E}[W_1]=1/\lambda<+\infty \qquad \text{and} \qquad \operatorname{Var}(W_1)=1/\lambda^2<+\infty,$$

leads to

$$\sqrt{n}\frac{\left(\frac{1}{\hat{\lambda}_n}-\frac{1}{\lambda}\right)}{\sqrt{1/\lambda^2}}=\sqrt{n}\frac{\left(\bar{W}_n-\frac{1}{\lambda}\right)}{\sqrt{1/\lambda^2}}=\sqrt{n}\frac{(\bar{W}_n-\mathbb{E}[W_1])}{\sqrt{\operatorname{Var}(W_1)}}\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}(0,1).$$

We deduce (3.1) by multiplying by $\sqrt{1/\lambda^2}$ (which multiplies the variance by $(1/\lambda^2)$).

2. Deduce Theorem 3.9 from the delta method.

Solution.

Let us apply the delta method to the differentiable function

$$f : x\in\mathbb{R}_+^*\mapsto\frac{1}{x}, \qquad \text{with, by (3.1),} \qquad \begin{cases} r_n=\sqrt{n},\\ X_n=1/\hat{\lambda}_n,\\ \theta=1/\lambda,\\ Y\sim\mathcal{N}(0,1/\lambda^2).\end{cases}$$

(a) Since $f(X_n)=\hat{\lambda}_n$ and $f(\theta)=\lambda$, we first obtain the LLN-type result:

$$\hat{\lambda}_n\xrightarrow[n\to+\infty]{\mathbb{P}}\lambda.$$

(b) Moreover, $\forall x\in\mathbb{R}_+^*$, $f'(x)=-1/x^2$. Hence,

$$f'(\theta)Y=-\lambda^2Y \quad\sim\quad \mathcal{N}\left(0,(-\lambda^2)^2\times[1/\lambda^2]\right),$$

and we obtain the CLT-type result:

$$\sqrt{n}\left(\hat{\lambda}_n-\lambda\right)\xrightarrow[n\to+\infty]{\mathcal{L}}\mathcal{N}\left(0,\lambda^2\right).$$

□

III.4 Confidence intervals and statistical tests

Fix n in $\mathbb{N}^*$ and let α in $(0,1)$. We observe the n first arrival times of a homogeneous Poisson process with intensity $\lambda>0$ and we aim at constructing confidence intervals for λ with confidence level $1-\alpha$, or at testing, for instance,

$$\mathcal{H}_0:\lambda\leq\lambda_0 \quad\text{against}\quad \mathcal{H}_1:\lambda>\lambda_0, \qquad \text{at level } \alpha.$$

A) Non-asymptotic statistical inference

Based on Proposition 1.18, one may construct a (non-asymptotic) confidence interval for λ with confidence level $1-\alpha$, that is

$$\left[\frac{x_{2n,\alpha/2}}{2T_n};\frac{x_{2n,1-\alpha/2}}{2T_n}\right],$$

where $x_{d,\eta}$ denotes the η-quantile of a $\chi^2(d)$ distribution.

Proof. In exercise.

Solution.

By Proposition 1.18, we know that $T_n \sim \Gamma(n,\lambda)$, thus, $2\lambda T_n \sim \Gamma(n,1/2) \overset{(d)}{=} \chi^2(2n)$.
In particular,

$$\mathbb{P}\left(\frac{x_{2n,\alpha/2}}{2T_n} \leq \lambda \leq \frac{x_{2n,1-\alpha/2}}{2T_n}\right) = \mathbb{P}\left(x_{2n,\alpha/2} \leq 2\lambda T_n \leq x_{2n,1-\alpha/2}\right) = 1-\alpha.$$

□

Equivalently, one may prove that the test which rejects $\mathcal{H}_0$ if $2\lambda_0 T_n \leq x_{2n,\alpha}$ is of (non-asymptotic) level α, that is

$$\sup_{\lambda\leq\lambda_0}\{\mathbb{P}_\lambda(2\lambda_0 T_n \leq x_{2n,\alpha})\} \leq \alpha.$$

The test statistic of this test is

$$X_n := 2\lambda_0 T_n.$$

Moreover, if $\lambda=\lambda_0$, $X_n \sim \chi^2(2n)$. Hence, the p-value of this test equals

$$\mathbb{P}_{\lambda_0}\left(X_n \leq X_n^{obs}\right) = F_{2n}\left(X_n^{obs}\right),$$

where X_n^{obs} is the observed value of the test statistic (numerical application) and F_{2n} denotes the c.d.f. of a $\chi^2(2n)$ distribution.

Proof. Facultative homework. □

B) Asymptotic statistical inference

From Theorem 3.9, one can construct an asymptotic confidence interval that is

$$\left[\frac{n}{T_n\left(1+\frac{z_{1-\alpha/2}}{\sqrt{n}}\right)};\frac{n}{T_n\left(1-\frac{z_{1-\alpha/2}}{\sqrt{n}}\right)}\right],$$

where $z_{1-\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of a standard gaussian distribution $\mathcal{N}(0,1)$.
Note that this confidence interval requires that $1-z_{1-\alpha/2}/\sqrt{n}>0$, that is $n>z^2_{1-\alpha/2}$, and does not need Slutsky's Lemma.

Proof. In exercise.

Solution.

From Theorem 3.9, we know that

$$\sqrt{n}\left(\frac{n}{T_n}-\lambda\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,\lambda^2).$$

Hence,

$$\sqrt{n}\left(\frac{n}{\lambda T_n}-1\right) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1).$$

We deduce that, if $1-\frac{z_{1-\alpha/2}}{\sqrt{n}}>0$,

$$\begin{aligned}\mathbb{P}\left(\frac{n}{T_n\left(1+\frac{z_{1-\alpha/2}}{\sqrt{n}}\right)}\le\lambda\le\frac{n}{T_n\left(1-\frac{z_{1-\alpha/2}}{\sqrt{n}}\right)}\right) &= \mathbb{P}\left(-z_{1-\alpha/2}\le\sqrt{n}\left(\frac{n}{\lambda T_n}-1\right)\le z_{1-\alpha/2}\right)\\ &\xrightarrow[T\to+\infty]{} \mathbb{P}\left(-z_{1-\alpha/2}\le Z\le z_{1-\alpha/2}\right)=1-\alpha,\end{aligned}$$

where Z denotes a standard gaussian r.v.

□

Moreover, one can prove that the test which rejects $\mathcal{H}_0$ if $\hat{\lambda}_n\ge\lambda_0(1+z_{1-\alpha}/\sqrt{n})$ is (uniformly) asymptotically of level α, that is

$$\limsup_{n\to+\infty}\left(\sup_{\lambda\le\lambda_0}\left\{\mathbb{P}_\lambda\left(\hat{\lambda}_n\ge\lambda_0\left(1+\frac{z_{1-\alpha}}{\sqrt{n}}\right)\right)\right\}\right)\le\alpha.$$

This result lies on Lemma 3.7.

Note that, by isolating the quantile in the rejection rule, we recover the test statistic, that is

$$Z_n:=\sqrt{n}\left(\frac{\hat{\lambda}_n}{\lambda_0}-1\right)=\sqrt{n}\left(\frac{n}{\lambda_0 T_n}-1\right).$$

Moreover, if $\lambda=\lambda_0$, $Z_n \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,1)$. Hence the p-value of this test equals

$$\mathbb{P}_{\lambda_0}\left(Z_n\ge Z_n^{obs}\right)\approx 1-\Phi\left(Z_n^{obs}\right),$$

where Z_n^{obs} denotes the observed value of the test statistic (numerical application), and Φ denotes the c.d.f. of a standard gaussian distribution $\mathcal{N}(0,1)$.

Proof. Facultative homework. □

Take home message

Estimator	Unbiased	Efficient	Asymptotically unbiased	Consistent	Asymptotically gaussian
$\hat{\lambda}_T = \frac{N_T}{T}$	✓	✓	✓	✓	$\sqrt{T}(\hat{\lambda}_T - \lambda) \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}(0,\lambda)$
$\hat{\lambda}_n = \frac{n}{T_n}$	✗	·	✓	✓	$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0,\lambda^2)$
$\tilde{\lambda}_n = \frac{n-1}{T_n}$	✓	✗	✓	✓	(?)

Chapter 4

Inhomogeneous Poisson processes

In this chapter, we consider inhomogeneous Poisson processes. They generalize homogeneous Poisson processes by allowing the rate of the process to be a varying function of time. In particular, by doing so, we drop the stationary increments property.

Inhomogeneous Poisson processes provide nice models in Reliability theory. For instance, if at each failure a system is repaired to its condition at the time of failure and placed in service again, then the failures are often modeled by an inhomogeneous Poisson process, provided the repair times can be neglected.

I Definitions and basic properties

Definition 4.1.

Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be a real valued piecewise continuous (or at least locally integrable) function defined on $\mathbb{R}_+$. A point process $(N_t)_{t\in\mathbb{R}_+}$ is said to be an inhomogeneous Poisson process with intensity function λ if

(1) $N_0 = 0$.

(2) The process has independent increments.

(3) For all $t > 0$, the number of points in $[0,t]$ has a Poisson distribution with parameter $\int_0^t \lambda(u)du$, that is

$$N_t \sim \mathcal{P}\left(\int_0^t \lambda(u)du\right).$$

The function $\Lambda : t \mapsto \int_0^t \lambda(u)du$ is called *cumulative intensity* of the Poisson process.

Example 4.2.

The power law process, sometimes called the Weibull process, is very popular in Reliability theory. It is an inhomogeneous Poisson process with intensity given by

$$\lambda(s) = \frac{\beta}{\alpha}\left(\frac{s}{\alpha}\right)^{\beta-1},$$

where α and β are two positive parameters.

Proposition 4.3.

Let N be an inhomogeneous Poisson process with intensity function λ.
Denote for all Borel subset A of $\mathbb{R}_+$, $N(A)$ the number of points in A. Note that if $(T_i)_{i\in\mathbb{N}^*}$

denote the arrival times, then,

$$N(A) = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{T_n \in A\}}$$

1. *For all bounded Borel set A in $\mathbb{R}_+$, $N(A)$ is a.s. finite.*
2. *For all $s \geq 0$ and all $t > 0$, the random variable $N(]s, s+t]) = N_{s+t} - N_s$ has a Poisson distribution with parameter $\mu(]s, s+t]) = \int_s^{s+t} \lambda(u)du$, that is*

$$N_{s+t} - N_s \sim \mathcal{P}\left(\int_s^{s+t} \lambda(u)du\right).$$

Proof. 1. Let A be a bounded Borel set in $\mathbb{R}_+$. Hence, there exists $t_A \in \mathbb{R}_+^*$ s.t. $A \subset [0, t_A]$. Therefore,

$$N(A) = \sum_{i \in \mathbb{N}^*} \mathbb{1}_{\{T_n \in A\}} \quad \leq \quad \sum_{i \in \mathbb{N}^*} \mathbb{1}_{\{T_n \leq t_A\}} = N_{t_A}.$$

Moreover, $N_{t_A} \sim \mathcal{P}(\int_0^{t_A} \lambda(u)du)$. Hence $\mathbb{E}[N_{t_A}] = \int_0^{t_A} \lambda(u)du < +\infty$. We deduce that, N_{t_A} is finite a.s., and thus, so is $N(A)$.

2. In exercise:

(a) Recall the moment generating function (or m.g.f.), or Laplace transform, of a Poisson r.v., and for all $t > 0$, express the m.g.f. of N_t w.r.t. the cumulative intensity Λ of the process.

Solution.

Let $X \sim \mathcal{P}(\mu)$. Then for all $u < 0$, $\mathbb{E}\left[e^{uX}\right] = \exp\left(\lambda(e^u - 1)\right)$.
Hence, for any $t > 0$, since $N_t \sim \mathcal{P}(\Lambda(t))$, then for all $u < 0$,

$$\mathbb{E}\left[e^{uN_t}\right] = \exp\left(\Lambda(t)\left[e^u - 1\right]\right).$$

(b) Let $s \geq 0$ and $t > 0$. Compute the m.g.f. of $N_{s+t} - N_s$, and conclude. *Hint:* Use $N_{s+t} = N_{s+t} - N_s + N_s$.

Solution.

Note that $N_{s+t} = N_{s+t} - N_s + N_s$, with $(N_{s+t} - N_s) \perp\!\!\!\perp N_s$ since the intervals $[0, s]$ and $(s, s+t]$ are disjoint. Hence

$$\mathbb{E}\left[e^{uN_{s+t}}\right] = \mathbb{E}\left[e^{u(N_{s+t}-N_s)}e^{uN_s}\right] = \mathbb{E}\left[e^{u(N_{s+t}-N_s)}\right]\mathbb{E}\left[e^{uN_s}\right].$$

Therefore, by the previous question,

$$\exp\left(\Lambda(t+s)\left[e^u - 1\right]\right) = \mathbb{E}\left[e^{u(N_{s+t}-N_s)}\right]\exp\left(\Lambda(s)\left[e^u - 1\right]\right),$$

i.e.

$$\mathbb{E}\left[e^{u(N_{s+t}-N_s)}\right] = \exp\left(\left[\Lambda(s+t) - \Lambda(s)\right]\left[e^u - 1\right]\right) = \exp\left(\int_s^{s+t} \lambda(x)dx\left[e^u - 1\right]\right).$$

We recognize the m.g.f. of a $\mathcal{P}\left(\int_s^{s+t} \lambda(u)du\right)$, which ends the proof.

□

Let us notice several points.

a) If for all $t \geq 0$, $\lambda(t) = \lambda$, that is λ is a constant function, $\int_s^{s+t} \lambda du = \lambda t$, and one recovers the definition of a homogeneous Poisson process with rate λ.

b) Notice that Definition 4.1 characterizes the distribution of the process N. Indeed, for any n and for any $0 \leq t_1 < t_2 < \ldots < t_n$, Definition 4.1 and Proposition 4.3 both provide the distribution of $(N_{t_1}, N_{t_2} - N_{t_1}, \ldots, N_{t_n} - N_{t_{n-1}})$. Hence, the distribution of the process $(N_t)_{t \in \mathbb{R}_+}$ is well known.

c) The terminology of point process theory has been criticized for being too varied. In addition to the word point often being omitted, the homogeneous Poisson (point) process is also called a stationary Poisson (point) process, as well as uniform Poisson (point) process. The inhomogeneous Poisson (point) process, as well as being called nonhomogeneous, is also referred to as the non-stationary Poisson process.

d) The positive measure μ defined on any Borel set A by $\mu(A) = \int_A \lambda(u)du$ (or equivalently $d\mu(u) = \lambda(u)du$) is called the *intensity measure* of the Poisson process N. This definition of an inhomogeneous Poisson process can be generalized to any non-negative intensity measure on $\mathbb{R}_+$ that is locally bounded (i.e. for all bounded Borel subset A in $\mathbb{R}_+$, $\mu(A) < +\infty$), not necessarily absolutely continuous with respect to the Lebesgue measure by replacing assumption (3) by for all Borel set A in $\mathbb{R}_+$, $N(A) \sim \mathcal{P}(\int_A d\mu)$.

   However, in this lecture, we restrict attention to the case $d\mu(u) = \lambda(u)du$, that is Poisson processes with intensity functions.

Property 4.4.

Let N be an inhomogeneous Poisson process with intensity function λ on $\mathbb{R}_+$.

1. The mean function of N equals

$$\forall t \geq 0, \quad m(t) = \mathbb{E}[N_t] = \int_0^t \lambda(u)du = \Lambda(t).$$

   In particular, the rate (see Definition 2.7) of the Poisson process N equals:

$$\forall t \geq 0, \quad w(t) = m'(t) = \lambda(t).$$

2. Moreover, for all $t \geq 0$, as $h \to 0$,

   (a) $\mathbb{P}(N_{t+h} - N_t = 0) = 1 - \lambda(t)h + o(h)$.
   (b) $\mathbb{P}(N_{t+h} - N_t = 1) = \lambda(t)h + o(h)$.
   (c) $\mathbb{P}(N_{t+h} - N_t \geq 2) = o(h)$.

Proof. Facultative exercise.

1. Immediate.

2. To do. *Hint:* Since $\Lambda'(t) = \lambda(t)$, one has $\Lambda(t+h) - \Lambda(t) = \lambda(t)h + o(h)$.

   Solution.

   By Property 4.4, $N_{t+h} - N_t \sim \mathcal{P}(\Lambda(t+h) - \Lambda(t))$. Then,

(a) $\mathbb{P}(N_{t+h} - N_t = 0) = e^{-[\Lambda(t+h)-\Lambda(t)]} = e^{-\lambda(t)h+o(h)} = 1 - \lambda(t)h + o(h)$.

(b)
$$\begin{aligned}\mathbb{P}(N_{t+h} - N_t = 1) &= [\Lambda(t+h) - \Lambda(t)]\, e^{-[\Lambda(t+h)-\Lambda(t)]} \\ &= [\lambda(t)h + o(h)] \times [1 - \lambda(t)h + o(h)] \\ &= \lambda(t)h + o(h).\end{aligned}$$

(c) $\mathbb{P}(N_{t+h} - N_t \geq 2) = 1 - [\mathbb{P}(N_{t+h} - N_t = 0) + \mathbb{P}(N_{t+h} - N_t = 1)]$.

□

As in the homogeneous case, one can provide an equivalent definition for inhomogeneous Poisson processes.

Definition 4.5.

The point process N is said to be an inhomogeneous Poisson process with intensity function λ defined on $\mathbb{R}_+$ if

(1) $N_0 = 0$.

(2) The process has independent increments.

(3) For all positive t, one has as $h \to 0$,

  (a) $\mathbb{P}(N_{t+h} - N_t = 1) = \lambda(t)h + o(h)$.
  (b) $\mathbb{P}(N_{t+h} - N_t \geq 2) = o(h)$.

The proof is the same as in the homogeneous case, except that we obtain a linear differential equation with non-constant coefficients.

Notice that, a Poisson process N with intensity function λ is regular since, for all $t > 0$,

$$\lim_{h \to 0} \frac{\mathbb{P}(N_{t+h} - N_t \geq 2)}{h} = 0.$$

In particular, it is a simple process, that is no more than 1 event can occur at the same time, and each jump of the counting process equals 1 (almost surely). Notice that it is not the case anymore when we generalize to inhomogeneous Poisson processes with non-diffuse intensity measure (i.e. that contains atoms).

II Construction of an inhomogeneous Poisson process

II.1 Time change

A first way of constructing an inhomogeneous Poisson process with intensity function λ on $\mathbb{R}_+$ (and thus proving that such processes exist) is by changing the time scale of a homogeneous Poison process. Before, let us recall the definition and some properties of the generalized inverse function.

A) Generalized inverse function

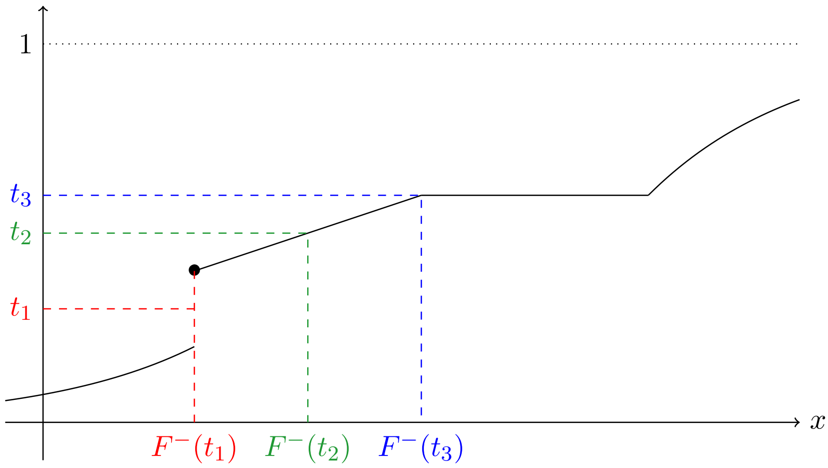
Definition 4.6.

Let $F : \mathbb{R} \to \mathbb{R}_+$ be a non-decreasing càdlàg non-negative function defined on $\mathbb{R}$. Its generalized inverse function is defined by

$$F^- : \begin{pmatrix} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \cup \{\pm\infty\} \\ t & \longmapsto & \inf\{x \in \mathbb{R} \;;\; F(x) \geq t\} \end{pmatrix},$$

with the conventions $\inf(\varnothing) = +\infty$ and $\inf(\mathbb{R}) = -\infty$.

Note that if F is a c.d.f., then F^- is the corresponding quantile function.



Property 4.7.

The generalized inverse function satisfies the following properties.

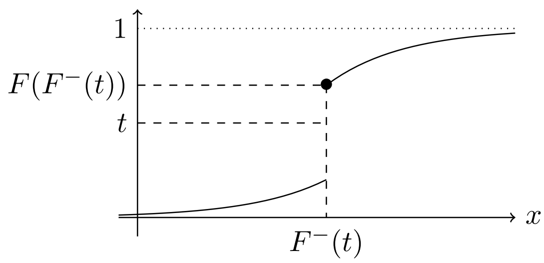
i. F^- is a non-decreasing function.

ii. $F(x) \geq t \quad \Leftrightarrow \quad x \geq F^-(t)$.

iii. F^- is left-continuous, that is for all t_0 in $\mathbb{R}$, $\lim\limits_{\substack{t \to t_0 \\ t < t_0}} F^-(t) = F^-(t_0)$.

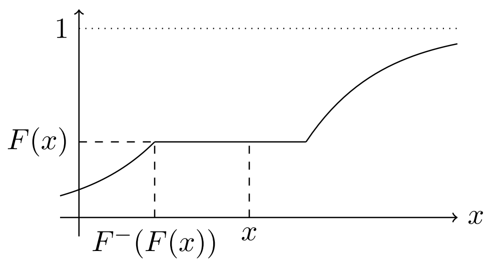
iv.

For all $t \in \mathbb{R}_+$, $F(F^-(t)) \geq t$.



Moreover, if F is continuous, for all $t \in \mathbb{R}_+$, $F(F^-(t)) = t$.

For all $x \in \mathbb{R}$, $F^-(F(x)) \leq x$.



Moreover, if F is strictly increasing, for all $x \in \mathbb{R}$, $F^-(F(x)) = x$.

v. In particular, if F is strictly increasing and continuous, then F is bijective with inverse function $F^{-1} = F^-$.

Proof. See Appendix A.II. □

B) Back to inhomogeneous Poisson processes

Proposition 4.8.

Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be a locally integrable function, and denote

$$\Lambda : x \in \mathbb{R}_+ \mapsto \int_0^x \lambda(u)du.$$

Note that Λ is a non-decreasing continuous function.

1. *Let $\tilde{N}$ be a homogeneous Poisson process with (constant) rate 1, and denote*

$$0 < S_1 < S_2 < \ldots < S_n < \ldots$$

its arrival times. Define respectively for all $n \in \mathbb{N}^*$, and for all $t \in \mathbb{R}_+$

$$T_n = \Lambda^-(S_n) \qquad \text{and} \qquad N_t = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{T_n \leq t\}},$$

where Λ^- denote the generalized inverse function of Λ. Then $N = (N_t)_{t \in \mathbb{R}_+}$ is an inhomogeneous Poisson process with intensity function λ.

2. *Conversely, assume that Λ is strictly increasing (and thus a bijective function, since it is always continuous). Let N be a inhomogeneous Poisson process with intensity function λ on $\mathbb{R}_+$. Denote*

$$0 < T_1 < T_2 < \ldots < T_n < \ldots,$$

its arrival times. Define respectively for all n in $\mathbb{N}^*$, and for all t in $\mathbb{R}_+$,

$$S_n = \Lambda(T_n) \qquad \text{and} \qquad \tilde{N}_t = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{S_n \leq t\}}.$$

Then the point process $\tilde{N} = (\tilde{N}_t)_{t \in \mathbb{R}_+}$ is a homogeneous Poisson process with (constant) rate 1.

Proof. In exercise.

1. (a) Let $t > 0$. Express N_t in terms of $(S_n)_{n \in \mathbb{N}^*}$, and deduce, by Property 4.7, that $N_t = \tilde{N}_{\Lambda(t)}$.

Solution.

$$N_t = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{T_n \leq t\}} = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{\Lambda^-(S_n) \leq t\}}.$$

Yet, by point ii. of Property 4.7,

$$\Lambda^-(S_n) \leq t \quad \Leftrightarrow \quad S_n \leq \Lambda(t).$$

Hence

$$N_t = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{S_n \leq \Lambda(t)\}} = \tilde{N}_{\Lambda(t)}.$$

(b) Conclude.

Solution.

i. $N_0 = \tilde{N}_0 = 0$ since $\tilde{N}$ is a homogeneous Poisson process and $\Lambda(0) = \int_0^0 \lambda(u)du = 0$.

ii. Let $t_1 < t_2 \le s_1 < s_2$. Then, since Λ is non-decreasing, then

$$\Lambda(t_1) \le \Lambda(t_2) \le \Lambda(s_1) \le \Lambda(s_2).$$

In particular, since $\tilde{N}$ is a homogeneous Poisson process,

$$N_{t_2} - N_{t_1} = \tilde{N}_{\Lambda(t_2)} - \tilde{N}_{\Lambda(t_1)} \quad \perp\!\!\!\perp \quad \tilde{N}_{\Lambda(s_2)} - \tilde{N}_{\Lambda(s_1)} = N_{s_2} - N_{s_1}.$$

Note that this remains true if $\Lambda(t_2) = \Lambda(t_1)$ (or if $\Lambda(s_2) = \Lambda(s_1)$) since in this case, the corresponding increment is constant (equal to zero).

iii. Finally, since $\tilde{N}$ is a homogeneous Poisson process with rate 1,

$$N_t = \tilde{N}_{\Lambda(t)} \sim \mathcal{P}\left(1 \times \Lambda(t)\right).$$

2. Assume that Λ is strictly increasing (and thus bijective), and denote Λ^{-1} its inverse function.

(a) What is the monotony of Λ^{-1}? Deduce that for all u, t,

$$\Lambda(u) \le t \quad \Leftrightarrow \quad u \le \Lambda^{-1}(t).$$

Solution.

Since Λ^{-1} is the inverse function of a strictly increasing function, it is also strictly increasing. We deduce that:

$\boxed{\Rightarrow}$ if $\Lambda(u) \le t$, then $u = \Lambda^{-1}(\Lambda(u)) \le \Lambda^{-1}(t)$,

$\boxed{\Rightarrow}$ if $u \le \Lambda^{-1}(t)$, then $\Lambda(u) \le \Lambda(\Lambda^{-1}(t)) = t$.

(b) Deduce that for all $t > 0$, $\tilde{N}_t = N_{\Lambda^{-1}(t)}$.

Solution.

Let $t > 0$. Then,

$$\tilde{N}_t = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{S_n \le t\}} = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{\Lambda(T_n) \le t\}} = \sum_{n \in \mathbb{N}^*} \mathbb{1}_{\{T_n \le \Lambda^{-1}(t)\}} = N_{\Lambda^{-1}(t)}.$$

(c) Conclude.

Solution.

i. Since $\Lambda(0) = 0$, then $\Lambda^{-1}(0) = 0$ and $\tilde{N}_0 = N_0 = 0$.

ii. $\tilde{N}$ has independent increments since N does.

iii. By Proposition 4.3,

$$\tilde{N}_{s+t} - \tilde{N}_s = N_{\Lambda^{-1}(t+s)} - N_{\Lambda^{-1}(s)} \sim \mathcal{P}\left(\int_{\Lambda^{-1}(s)}^{\Lambda^{-1}(s+t)} \lambda(u)du\right),$$

with $\displaystyle\int_{\Lambda^{-1}(s)}^{\Lambda^{-1}(s+t)} \lambda(u)du = \Lambda\left(\Lambda^{-1}(s+t)\right) - \Lambda\left(\Lambda^{-1}(s)\right) = s+t-s = t$. Hence,

$$\tilde{N}_{s+t} - \tilde{N}_s \sim \mathcal{P}(1 \times t).$$

We deduce from all three points above that $\tilde{N}$ is a homogeneous Poisson process with (constant) rate 1.

□

II.2 Acceptance/Rejection (or Thinning)

The following result leads to one of the most popular method for generating an inhomogeneous Poisson process, sometimes called the *thinning method*. It is the "point process" analogue of the acceptance/rejection method classically used to simulate r.v.

It generalizes Proposition 2.27 which claims that if you randomly classify each point of a homogeneous Poisson process with rate λ into two categories (with probability p and $1-p$), the two corresponding point processes are also homogeneous Poisson processes with corresponding rates $p\lambda$ and $(1-p)\lambda$. In the inhomogeneous case, we allow p to be a varying function of time.

Theorem 4.9.
Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+^*$ be a locally integrable non-negative function. Assume that λ is upper-bounded by $M \in \mathbb{R}_+^*$ on a subset $[0,T)$ of $\mathbb{R}_+$ ($T \in \mathbb{R}_+^* \cup \{+\infty\}$), that is

$$\forall u \in [0,T), \quad \lambda(u) \leq M.$$

Let N be a homogeneous Poisson process with rate M and classify each point u of N as type I with probability $p(u) := \lambda(u)/M$ and type II otherwise. Then, the resulting point processes N^{I} and N^{II} of type I and type II events satisfy

a) N^{I} is a inhomogeneous Poisson process with intensity function λ.
N^{II} is a inhomogeneous Poisson process with intensity function $M-\lambda$.

b) $N^{\mathrm{I}} \perp\!\!\!\perp N^{\mathrm{II}}$.

Idea of proof. a) The independence of the increments comes from the independence of the increments of the homogeneous Poisson process N, and the independence between the tags.

b) Let us now prove the Poisson distribution, and that $N^{\mathrm{I}} \perp\!\!\!\perp N^{\mathrm{II}}$.

- **Modeling:** We know (c.f. Proposition 2.23) that conditionally on $\{N_t = n\}$, $T_1, \ldots, T_n$ behaves as the order statistic $U_{(1)} \leq \ldots \leq U_{(n)}$ associated to n i.i.d. r.v. $U_1, \ldots, U_n$ with distribution $\mathcal{U}([0,T])$. The classifying process can be seen as follows: Let $Y_1, \ldots, Y_n$ be n i.i.d. r.v. with distribution $\mathcal{U}([0,M])$, independent on $(U_1, \ldots, U_n)$.

$$\text{Classify } U_i \text{ as} \quad \begin{cases} \text{type I if } Y_i \leq \lambda(U_i), \\ \text{type II otherwise.} \end{cases}$$

Then,

$$\begin{aligned}
\mathbb{P}(U_i \text{ of type I}|U_i = u) &= \mathbb{P}(Y_i \leq \lambda(U_i)|U_i = u) \\
&= \mathbb{P}(Y_i \leq \lambda(u)|U_i = u) \\
&= \mathbb{P}(Y_i \leq \lambda(u)), && \text{as } Y_i \perp\!\!\!\perp U_i \\
&= \frac{\lambda(u)}{M} = p(u) && \text{as } Y_i \sim \mathcal{U}([0, M]).
\end{aligned}$$

- **Towards a binomial distribution:** Conditionally on $\{N_t = n\}$, we thus obtain

$$N_t^{\mathrm{I}} = \sum_{i=1}^{n} \mathbb{1}_{\{Y_i \leq \lambda(U_i)\}},$$

where the $\mathbb{1}_{\{Y_i \leq \lambda(U_i)\}}$ are independent (by independence of the couples $(U_i, Y_i)_i$) Bernoulli r.v. with parameter

$$\tilde{p} := \quad \mathbb{P}(Y_i \leq \lambda(U_i)) = \int_0^{+\infty} \underbrace{\mathbb{P}(Y_i \leq \lambda(U_i)|U_i = u)}_{p(u)} \; \underbrace{f_{U_i}(u)}_{\frac{1}{t}\mathbb{1}_{\{u \in [0,t]\}}} \; du = \frac{1}{t}\int_0^t p(u)du.$$

We deduce that the conditional distribution of N_t^{I} given $\{N_t = n\}$, is a binomial distribution with parameters n and $\tilde{p}$, that is

$$\mathcal{L}\left(N_t^{\mathrm{I}} \middle| N_t = n\right) = \mathcal{B}(n, \tilde{p}).$$

- **Unconditioning:** Let $k, l \in \mathbb{N}$. Then

$$\begin{aligned}
\mathbb{P}\big(N_t^{\mathrm{I}} = k, N_t^{\mathrm{II}} = l\big) &= \mathbb{P}\big(N_t^{\mathrm{I}} = k, N_t = k + l\big) \\
&= \mathbb{P}\big(N_t^{\mathrm{I}} = k \big| N_t = k + l\big)\, \mathbb{P}(N_t = k + l) \\
&= \binom{k+l}{k} \tilde{p}^k (1 - \tilde{p})^l \; \times \; \frac{(Mt)^{k+l}}{(k+l)!} e^{-Mt(\tilde{p}+1-\tilde{p})} \\
&= \left[\frac{(Mt\tilde{p})^k}{k!} e^{-Mt\tilde{p}}\right] \times \left[\frac{(Mt[1-\tilde{p}])^l}{l!} e^{-Mt(1-\tilde{p})}\right].
\end{aligned}$$

We deduce that

$$\mathbb{P}\big(N_t^{\mathrm{I}} = k\big) = \sum_{l=0}^{+\infty} \mathbb{P}\big(N_t^{\mathrm{I}} = k, N_t^{\mathrm{II}} = l\big) = \frac{(Mt\tilde{p})^k}{k!} e^{-Mt\tilde{p}},$$

thus

$$N_t^{\mathrm{I}} \sim \mathcal{P}(Mt\tilde{p}), \qquad \text{with} \qquad Mt\tilde{p} = Mt\frac{1}{t}\int_0^t \frac{\lambda(u)}{M} du = \int_0^t \lambda(u)du.$$

In the same way, we prove that $N_t^{\mathrm{II}} \sim \mathcal{P}(Mt[1 - \tilde{p}])$, with

$$Mt[1 - \tilde{p}] = Mt - \int_0^t \lambda(u)du = \int_0^t [M - \lambda(u)]\, du.$$

Finally, we also deduce that $N_t^{\mathrm{I}} \perp\!\!\!\perp N_t^{\mathrm{II}}$ because of that product form above.

□

Algorithm 1 *Thinning algorithm*

Input: intensity function λ; simulation interval $[0, T]$; upper-bound $M > 0$.

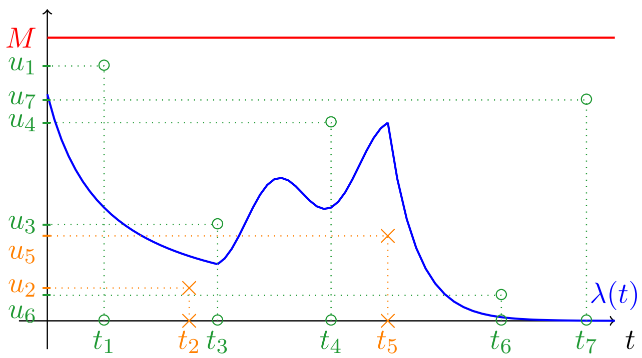
1. Generate a homogeneous Poisson process with constant rate M on $[0, T]$;
   Denote $t_1, \dots, t_n$ the observed arrival times on $[0, T]$.

2. Generate independently n i.i.d r.v. with distribution $\mathcal{U}[0, M]$;
   Denote $u_1, \dots, u_n$ the observed values.

3. for i in 1 to n, keep t_i if $u_i \leq \lambda(t_i)$.

4. Order the kept points.

Output: the set of ordered kept points are the arrival times of an inhomogeneous Poisson process with intensity function λ on $[0, T]$.



Observed arrival times:

- N: $t_1, \dots, t_7$;
- N^{I}: t_2, t_5;
- N^{II}: t_1, t_3, t_4, t_6, t_7.

III Arrival times

III.1 First arrival time after a given instant

Property 4.10.

Let N be a Poisson process with intensity function $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$.

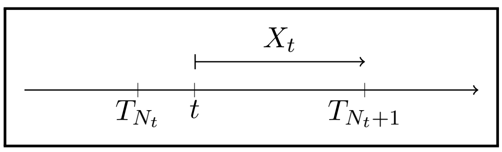
i. Then the survival function of the first arrival time T_1 equals for all $s > 0$,

$$\mathbb{P}(T_1 > s) = \mathbb{P}(N_s = 0) = e^{-\int_0^s \lambda(u)du} = e^{-\Lambda(s)}.$$

It is not exponentially distributed anymore (except if $\Lambda(s) = \lambda s$, i.e. if the Poisson process is homogeneous). Note that λ is the hazard rate of T_1.

ii. Denote X_t the waiting time at time t, that is

$$X_t = T_{N_t+1} - t.$$



Then, for all $s > 0$,

$$\mathbb{P}(X_t > s) = \mathbb{P}(N_{t+s} - N_t = 0) = e^{-\int_t^{s+t} \lambda(u)du} = e^{-\int_0^s \lambda(u+t)du}.$$

In particular, from Proposition 1.14, the hazard rate of X_t is

$$h_t : u \mapsto \lambda(u+t).$$

III.2 (Conditional) distribution of the arrival times

Proposition 4.11.

Let N be an inhomogeneous Poisson process with intensity function $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$. Fix $n \in \mathbb{N}^*$ and denote $(T_1, \ldots, T_n)$ the n first arrival times. Then $(T_1, \ldots, T_n)$ has the following joint density w.r.t. the Lebesgue mesure on $\mathbb{R}^n$:

$$(t_1, \ldots, t_n) \mapsto e^{-\int_0^{t_n} \lambda(u)du} \left(\prod_{i=1}^n \lambda(t_i) \right) \mathbb{1}_{\{0<t_1<\ldots<t_n\}}.$$

Note that if λ is a constant function, we recover the density in the continuous case (c.f. Lemma 2.19).

Proof. Admitted here.
Main argument: mathematical induction on n, using the Time change Proposition 4.8. □

We can thus deduce the conditional distribution of the arrival times.

Proposition 4.12.

Let $n \in \mathbb{N}^*$ and $t \in \mathbb{R}_+^*$. Given that $\{N_t = n\}$, the n arrival times $0 < T_1 < \ldots < T_n$ in $[0,t]$ of an inhomogeneous Poisson process with intensity function $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$, has conditional density w.r.t. the Lebesgue measure on $\mathbb{R}^n$:

$$(t_1, \ldots, t_n) \mapsto \frac{n!}{\int_0^t \lambda(u)du} \left(\prod_{i=1}^n \lambda(t_i) \right) \mathbb{1}_{\{0<t_1<\ldots<t_n\leq t\}}.$$

In particular, conditionally on $\{N_t = n\}$, $(T_1, \ldots, T_n)$ behaves as the order statistic associated to n i.i.d. r.v. with common density

$$s \in \mathbb{R} \mapsto \frac{\lambda(s)}{\int_0^t \lambda(u)du} \mathbb{1}_{\{0<s\leq t\}}.$$

Proof. Similar to the one of Proposition 2.23. □

This result leads to a new construction of an inhomogeneous Poisson process.

Proposition 4.13.

Let $\lambda : \mathbb{R}_+ \to \mathbb{R}_+$ be an intensity function. Let $Y \sim \mathcal{P}\left(\int_0^T \lambda(u)du\right)$.

- *If $Y = 0$, then for all $0 \leq t \leq T$, $N_t = 0$ (the point process has no points in $[0,T]$.*
- *Otherwise, let $S_1, \ldots, S_Y$ be i.i.d. r.v. with density w.r.t. the Lebesgue measure on $\mathbb{R}$*

$$s \in \mathbb{R} \mapsto \frac{\lambda(s)}{\int_0^T \lambda(u)du} \mathbb{1}_{\{0<s\leq T\}}.$$

Sort the values $(T_1, \ldots, T_Y) = (S_{(1)}, \ldots, S_{(Y)})$ and consider the associated counting pro-

cess defined for all $0 \leq t \leq T$,

$$N_t = \sum_{n=1}^{Y} \mathbb{1}_{\{T_n \leq t\}} = \sum_{i=1}^{Y} \mathbb{1}_{\{S_i \leq t\}}.$$

Then $(N_t)_{t\in[0,T]}$ is an inhomogeneous Poisson process with intensity function λ on $[0,T]$.

Proof. • On the one hand, on the event $\{Y=0\}$, $N_t = 0$.

• On the other hand, on the event $\{Y=n\}$,

$$N_t = \sum_{i=1}^{n} \mathbb{1}_{\{S_i \leq t\}}.$$

To prove that $(N_t)_{t\in[0,T]}$ is a Poisson process, we need to prove that for all subdivision of $[0,T]$,

$$0 = t_0 < t_1 < \ldots < t_{m-1} < t_m = T,$$

and for all integers $k_1, \ldots, k_m$, one has

$$\mathbb{P}\big(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m\big) = \prod_{j=1}^{m} \left(\frac{c_j^{k_j}}{k_j!} e^{-c_j} \right), \tag{4.1}$$

where

$$c_j = \int_{t_{j-1}}^{t_j} \lambda(u)du = \Lambda(t_j) - \Lambda(t_{j-1}).$$

Yet, since $N_{t_m} = N_T = Y$ (there are Y S_i's that belong to $[0,T]$),

$$\begin{aligned}
\mathbb{P}\big(N_{t_1} = k_1, &N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m\big) \\
&= \mathbb{P}\left(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m, Y = \sum_{j=1}^{m} k_j \right) \\
&= \mathbb{P}\left(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m \middle| Y = \sum_{j=1}^{m} k_j \right) \mathbb{P}\left(Y = \sum_{j=1}^{m} k_j \right)
\end{aligned}$$

On the one hand, for all j, for all i, by definition of the S_i's,

$$\mathbb{P}(S_i \in (t_{j-1}, t_j]) = \int_{t_{j-1}}^{t_j} \frac{\lambda(s)}{\Lambda(T)} ds = \frac{c_j}{c},$$

where c denotes

$$c = \int_0^T \lambda(u)du = c_1 + \ldots + c_m = \Lambda(T).$$

Hence, conditionally on $\left\{Y = \sum_{j=1}^m k_j\right\}$, the random variables $\left(\sum_{i=1}^{k_1+\ldots+k_m} \mathbb{1}_{\{S_i\in(t_{j-1},t_j]\}}\right)_{1\leq j\leq m}$, which count the number of S_i's in each interval $(t_{j-1}, t_j]$, have a multinomial distribution with parameters

$(k_1 + \ldots + k_m)$ and $(c_j/c)_{1\le j\le m}$. Therefore,

$$\begin{aligned}
\mathbb{P}\left(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m \middle| Y = \sum_{j=1}^{m} k_j \right)
&= \mathbb{P}\left(\sum_{i=1}^{k_1+\ldots+k_m} \mathbb{1}_{\{S_i \in (t_0,t_1]\}} = k_1, \ldots, \sum_{i=1}^{k_1+\ldots+k_m} \mathbb{1}_{\{S_i \in (t_{m-1},t_m]\}} = k_m \right) \\
&= \frac{(k_1+k_2+\ldots+k_m)!}{k_1!\ldots k_m!} \left(\frac{c_1}{c}\right)^{k_1} \ldots \left(\frac{c_m}{c}\right)^{k_m}
\end{aligned}$$

On the other hand, $Y \sim \mathcal{P}(\Lambda(T))$, with $\Lambda(T) = c = c_1 + \ldots + c_m$.
Finally,

$$\begin{aligned}
\mathbb{P}\big(N_{t_1} = k_1, N_{t_2} - N_{t_1} = k_2, \ldots, N_{t_m} - N_{t_{m-1}} = k_m\big)
&= \frac{(k_1+k_2+\ldots+k_m)!}{k_1!\ldots k_m!} \left(\frac{c_1}{c}\right)^{k_1} \ldots \left(\frac{c_m}{c}\right)^{k_m} \frac{c^{\sum_{j=1}^m k_j}}{(k_1+\ldots+k_m)!} e^{-c} \\
&= \frac{c_1^{k_1}}{k_1!} e^{-c_1} \times \ldots \times \frac{c_m^{k_m}}{k_m!} e^{-c_m},
\end{aligned}$$

which ends the proof of (4.1). □

IV Parametric statistics for inhomogeneous Poisson processes

Let us observe an inhomogeneous Poisson process N with intensity function λ_θ which depends on an unknown parameter $\theta \in \Theta$. Then, one can introduce the Maximum Likelihood Estimator (MLE) of θ. To do so, we need to compute the likelihood $\mathcal{L}(N;\theta)$, and, if it exists, to compute the MLE of θ,

$$\hat{\theta}_{MLE} \in \underset{\theta\in\Theta}{\operatorname{argmax}} \mathcal{L}\left(N;\theta\right),$$

and estimate the intensity function by $\lambda_{\hat{\theta}_{MLE}}$.

As in the homogeneous case, there are two types of observation: either the observation window is fixed (and you observe a random number of events), or n is fixed (and the observation window is random).

Proposition 4.14 (Likelihood in the "fixed window case").
Let N be an inhomogeneous Poisson process with intensity function $\lambda_\theta : \mathbb{R}_+ \to \mathbb{R}_+$ depending on an unknown parameter θ in Θ. Consider the number N_T of "events" that have occurred in $[0,T]$, and denote $(T_1, \ldots, T_{N_T})$ the corresponding arrival times. Then likelihood of θ is equal to

$$\mathcal{L}\left((N_t)_{t\in[0,T]};\theta\right) = \left(\prod_{i=1}^{N_T} \lambda_\theta(T_i)\right) e^{-\int_0^T \lambda_\theta(u)du}.$$

Proposition 4.15 (Likelihood in the "fixed number of points case").
Let N be an inhomogeneous Poisson process with intensity function $\lambda_\theta : \mathbb{R}_+ \to \mathbb{R}_+$ depending on an unknown parameter θ in Θ. Fix $n \in \mathbb{N}^*$ and denote $(T_1, \ldots, T_n)$ the n first arrival times. Then likelihood of θ is equal to

$$\mathcal{L}\left((T_i)_{1\le i\le n};\theta\right) = \left(\prod_{i=1}^{n} \lambda_\theta(T_i)\right) e^{-\int_0^{T_n} \lambda_\theta(u)du}.$$

You will have the opportunity to deepen your understanding of statistical inference for inhomogeneous Poisson processes through various mini-projects.

Appendix A

Mathematical tools

I Integration by substitution

Theorem A.1.

Let U and V be two open sets in $\mathbb{R}^n$ and $\phi : U \to V$ a C^1-diffeomorphism. Then, for any real-valued continuous and integrable function f defined on V,

$$\int_{\phi(U)} f(x)dx = \int_U f(\phi(u))|\det(J_\phi(u))|du.$$

where $J_\phi(u)$ denotes the Jacobian matrix of $\phi = (\phi_1, \ldots, \phi_n)$ at point $u = (u_1, \ldots, u_n)$, that is

$$J_\phi(u) = \begin{pmatrix} \frac{\partial \phi_1}{\partial u_1} & \cdots & \frac{\partial \phi_1}{\partial u_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial u_1} & \cdots & \frac{\partial \phi_n}{\partial u_n} \end{pmatrix}.$$

II Generalized inverse function

Definition A.2.

Let $F : \mathbb{R} \to \mathbb{R}_+$ be a non-decreasing càdlàg non-negative function defined on $\mathbb{R}$. Its generalized inverse function is defined by

$$F^- : \begin{pmatrix} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \cup \{\pm\infty\} \\ t & \longmapsto & \inf\{x \in \mathbb{R} \,;\, F(x) \geq t\} \end{pmatrix},$$

with the conventions $\inf(\varnothing) = +\infty$ and $\inf(\mathbb{R}) = -\infty$.

Note that if F is a c.d.f., then F^- is the corresponding quantile function.

Property A.3.

The generalized inverse function satisfies the following properties.

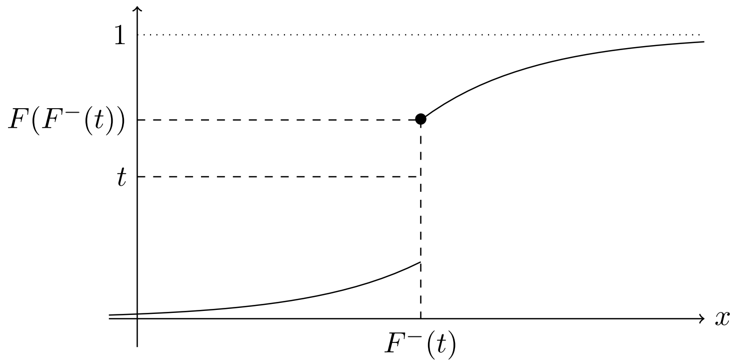
i. F^- is a non-decreasing function.

ii. $F(x) \geq t \quad \Leftrightarrow \quad x \geq F^-(t)$.

iii. F^- is left-continuous, that is

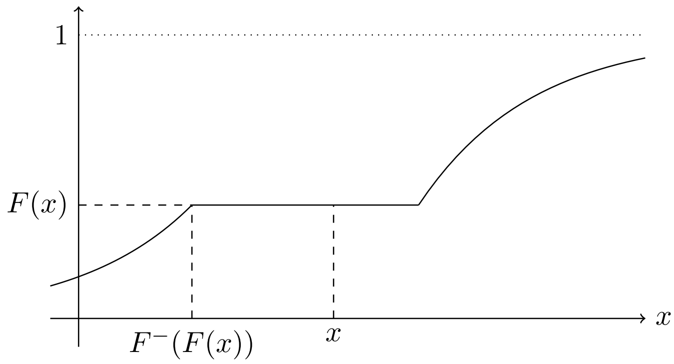
$$\lim_{\substack{t\to t_0\\ t<t_0}} F^-(t) = F(t_0).$$

iv. For all $t \in \mathbb{R}_+$, $F(F^-(t)) \geq t$.



Moreover, if F is continuous, for all $t \in \mathbb{R}_+$, $F(F^-(t)) = t$.

v. For all $x \in \mathbb{R}$, $F^-(F(x)) \leq x$.



Moreover, if F is strictly increasing, for all $x \in \mathbb{R}$, $F^-(F(x)) = x$.

vi. In particular, if F is strictly increasing and continuous, then F is bijective with inverse function $F^{-1} = F^-$.

Proof. Denote for all $t \in \mathbb{R}^+$,

$$A_t = \{x \in \mathbb{R}\ ;\ F(x) \geq t\},$$

such that $F^-(t) = \inf(A_t)$.

i. If $s \leq t$, then, for all $x \in \mathbb{R}$, $F(x) \geq t \Rightarrow F(x) \geq s$. Hence, $A_t \subset A_s$ and thus,

$$F^-(t) = \inf(A_t) \geq \inf(A_s) = F^-(s).$$

ii. $\boxed{\Rightarrow}$ If $F(x) \geq t$, then $x \in A_t$ and thus $F^-(t) = \inf(A_t) \leq x$.

$\boxed{\Leftarrow}$ Assume $x \geq F^-(t)$. Let $\varepsilon > 0$. Then $x + \varepsilon > x \geq F^-(t)$, so $x + \varepsilon > \inf(A_t)$. In particular, there exists $x_\varepsilon \in A_t$ such that $x_\varepsilon \leq x + \varepsilon$. Since F is non-decreasing, then $F(x + \varepsilon) \geq F(x_\varepsilon) \geq t$, as $x_\varepsilon \in A_t$.

We proved that, for all $\varepsilon > 0$, $F(x+\varepsilon) \geq t$. Hence, since F is right-continuous, as $\varepsilon \to 0^+$,

$$F(x) = \lim_{\substack{\varepsilon \to 0 \\ \varepsilon > 0}} F(x+\varepsilon) \geq t.$$

iii. Admitted here. Note that F^- is not right-continuous as soon as F is not strictly increasing.

iv. By applying ii. to $x = F^-(t)$ directly leads to $F(F^-(t)) = F(x) \geq t$.

Assume now that F is continuous. Consider $(x_n)_n$ a sequence satisfying

$$\forall n, \quad x_n < F^-(t) \qquad \text{and} \qquad \lim_{n \to +\infty} x_n = F^-(t).$$

Then by contraposition of ii., for all n, $F(x_n) < t$, and by taking $n \to +\infty$, as F is continuous (and in particular left-continuous),

$$F(F^-(t) = \lim_{n \to +\infty} F(x_n) \leq t.$$

Finally, in this case, both $F(F^-(t)) \leq t$ and $F(F^-(t)) \geq t$ imply that $F(F^-(t)) = t$.

v. By applying ii. to $t = F(x)$ leads to $F^-(F(x)) = F^-(t) \leq x$.

Assume know that F is strictly increasing. Assume (absurd) that $F^-(F(x)) < x$. Then

$$F(F^-(F(x))) < F(x),$$

which is impossible by the first point applied to $t = F(x)$.

vi. Immediate from the previous point.

$\square$

Appendix B

Tutorials and Computer Lab

Worksheet 1 - Some reminders in probability

Exercise 1. Memoryless property and Defective cellphones.

(A) Generalization of the memoryless property.
Let X be an exponentially distributed r.v. $X \sim \mathcal{E}(\lambda)$, with parameter $\lambda > 0$, and U be a continuous real-valued random variable with density g. Assume $X \perp\!\!\!\perp U$. Prove that for all $t, s > 0$,

$$\mathbb{P}(X + U > t + s | X > t) = \mathbb{P}(X + U > s).$$

Hint: One may first prove that

$$\mathbb{P}(X + U > t + s | X > t) = \frac{\mathbb{P}(\{X > t + s - U\} \cap \{s - U > 0\})}{\mathbb{P}(X > t)} + \frac{\mathbb{P}(\{X > t\} \cap \{s - U \leq 0\})}{\mathbb{P}(X > t)}$$

(B) Cellphone warranty. A cellphone manufacturer offers warranty for 8 months with a free replacement if the device is defective. The manufacturer assumes that each unit has a lifetime that can be modeled by an exponential r.v. X with mean equal to 24 months.

1. What is the (theoretical) proportion of defective devices that will be replaced for free?
2. What warranty should the manufacturer offer in order to replace only 10% of the devices for free?
3. Compute the (conditional) distribution function define by

   $$H : t \mapsto \mathbb{P}(X \leq t | X \geq 8).$$

   Deduce that the density function for duration of devices that do not have any failure before 8 months equals

   $$h : t \mapsto \lambda e^{-\lambda(t-c)} \mathbb{1}_{\{t \geq c\}},$$

   where c, λ are parameters to determine. Note that this distribution is called "translated exponential distribution".
4. Charles is a VIP member, he has infinite warranty. He can replace his phone every time it fails. The manufacturer would like to know the probability that 3 phones will not cover 5 years if Charles' first cellphone lasts more than a year. Compute this conditional probability assuming the lifetimes of Charles' cellphones are i.i.d. exponential r.v.
   Hint: One may use the generalized memoryless property.

Exercise 2. Car batteries.

The lifetime (in years) of an automobile battery is described by a r.v. X having an exponential distribution with parameter $\lambda = 1/5$.

1. Determine the expected lifetime of the battery and its standard deviation.
2. Compute the probability that the battery lasts for at least 10 years.
3. If the battery has lasted for 5 years, what is the probability that it will last for at least an additional 5 years ?

Exercise 3. The Gamma distribution. Let $\alpha, \lambda > 0$ and consider $X \sim \Gamma(\alpha, \lambda)$.

1. Compute the Laplace transform of X.
2. Deduce $\mathbb{E}[X]$ and $\mathrm{Var}(X)$.
3. Let $\beta > 0$ and consider $Y \sim \Gamma(\beta, \lambda)$ independent of X. Show that $X + Y \sim \Gamma(\alpha + \beta, \lambda)$.
4. Deduce the distribution of the sum of n i.i.d. exponential random variables with parameter λ.

Exercise 4. The Weibull distribution. With the exponential distribution, the Weibull distribution is one of the most widely used lifetime distribution in reliability. Let $\alpha, \beta > 0$. A non-negative real-valued random variable X with density (w.r.t. the Lebesgue measure)

$$x \mapsto \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\alpha}\right)^{\beta}\right) \mathbb{1}_{\{x>0\}},$$

is said to have a Weibull distribution with parameter (α, β). We denote $X \sim \mathcal{W}(\alpha, \beta)$. Note that if $\beta = 1$, we recover an exponential density with parameter $1/\alpha$.

1. Prove that the survival function of X is $G : t \in \mathbb{R}_+ \mapsto \exp\left(-\left(\dfrac{t}{\alpha}\right)^{\beta}\right)$.
2. Deduce that the hazard rate of X equals $h(t) = \dfrac{\beta}{\alpha}\left(\dfrac{t}{\alpha}\right)^{\beta-1}$.
3. Prove that $\left(\dfrac{X}{\alpha}\right)^{\beta}$ has an exponential distribution with parameter 1.
4. Show that if $X_1, \ldots, X_n$ are n independent r.v. with distribution $\mathcal{W}(\alpha, \beta)$, then

$$\min_{1 \le i \le n} X_i \sim \mathcal{W}\left(\left(\frac{\alpha}{n^{1/\beta}}\right), \beta\right).$$

Worksheet 2 - Homogeneous Poisson process

Exercise 1. Defective cellphones (point process approach).

Consider Part **(B) "Cellphone warranty"** of Exercise 1, Worksheet 1.
A cellphone manufacturer offers warranty for 8 months with a free replacement if the device is defective. The manufacturer assumes that each unit has a lifetime that can be modeled by an exponential variable with mean equal to 24 months.

Recall that Charles is a VIP member, he has infinite warranty. He can replace his phone every time it fails. The manufacturer would like to know the probability that 3 phones will not cover 5 years if Charles' first cellphone lasts more than a year. Compute this conditional probability assuming the number of times Charles replaces his cellphone can be modeled by a homogeneous Poisson process.

Exercise 2. The inspection paradox (or *Le paradoxe de l'autobus* in French).

You have an infinite stock of light bulbs with lifetimes which can be modeled by i.i.d. random variables with exponential distribution with parameter $\lambda > 0$. At time $t = 0$, we switch on the light and we change the light bulb as soon as it burns out. The times we change the light bulbs

$$0 < T_1 < T_2 < \ldots < T_n < \ldots$$

thus define a Poisson process $(N_t)_{t\geq 0}$ with rate λ, where N_t denotes the number of consumed light bulbs by time t. We decide to approximate the parameter λ by the following "inspection" method:

- Fix a time $t > 0$.
- Observe the lifetime of the "inspected" light bulb at time t, that is $T_{N_t+1} - T_{N_t}$.

We denote $X_t = T_{N_t+1} - t$ the residual lifetime of the inspected bulb at time t, and $Y_t = t - T_{N_t}$ the "age" of the inspected bulb at time t.

1. Prove that $X_t \sim \mathcal{E}(\lambda)$.
2. Compute the survival function of Y_t. Is it exponentially distributed?
3. Prove that X_t and Y_t are independent random variables.
4. Compute the average lifetime of the inspected bulb at time t.
5. Compare it to the average lifetime of the first bulb.

We can see that, for all time $t > 0$, the expected lifetime of the inspected bulb is greater than the average lifetime of the non-inspected bulbs. This phenomenon is often referred to the inspection paradox.

Exercise 3. Small/Large claims. The time unit considered here is a week. A small insurance company has in average one claim per day. It classifies each claim in two categories;

- *small claim* if the claim is less than or equal to 5000 €,
- *large claim* if the claim is greater than 5000 €.

It estimates that large claims occur in 10% of the cases. We model the claim occurrences by a homogeneous Poisson process N with rate $\lambda > 0$, and denote N^I (resp. N^{II}) the point process counting the small claims (resp. large claims).

1. Determine the rate λ.
2. What is the probability of at least 1 large claim in a week?
3. Knowing there has been 5 large claims in 2 weeks, what is the expected number of claims in that interval?
4. There has been 30 accidents in February. What is the probability that 3 of them where large claims?

Worksheet 3 - Statistics for homogeneous Poisson process

Exercise 1. Unbiased estimation for a fixed number of observations.
In this exercice, we assume that $n \geq 2$ is fixed, and we observe a homogeneous Poisson process with rate $\lambda > 0$, up to the nth event. Denote $0 < T_1 < T_2 < \ldots < T_n$ the corresponding arrival times.

We estimate the unknown parameter λ by the maximum likelihood estimator (MLE) $\hat{\lambda}_n = n/T_n$.

1. Prove that the MLE is a biased estimator of the rate λ.
2. Consider $\tilde{\lambda}_n = (n-1)/T_n$.
   (a) Check that $\tilde{\lambda}_n$ is an unbiased estimator of λ.
   (b) Is it efficient?

Exercise 2. Quality control.
A manufacturing plant wants to check the quality of its production. To do so, the person in charge keeps count of the number of defective parts.

He thus denotes N_t the number of defective parts manufactured by time t (for any t in hours) and assumes $(N_t)_{t\in\mathbb{R}_+}$ is a Poisson process with unknown intensity $\lambda > 0$. He believes that the quality of the production is not acceptable if the average number of defective parts per hour is greater than or equal to 5.

1. (a) Do you agree with the Poisson process assumption?
   (b) We aim at testing
   $$\mathcal{H}_0 : \lambda \geq \lambda_0 \quad \text{against} \quad \mathcal{H}_1 : \lambda < \lambda_0.$$
   Explicit the value of λ_0 and justify the choice of these testing hypotheses.
2. The person in charge fixes the inspection time to one day and observes $(N_t)_{t\in[0,24]}$. By the end of the day, he counts 108 defective parts.
   (a) Determine an estimation of the rate λ.
   (b) According to you, is the production quality satisfactory?
3. Now, the person in charge decides to observe up to the hundredth flaw and notices that it appears after 22 hours and 15 minutes of production.
   (a) What is an estimation of λ in this case?
   (b) First, we may consider that 100 points is large enough to work with the asymptotic distribution. Given this new observation, is the production quality satisfactory?
   (c) Second, we do not consider the asymptotic distribution anymore. What is your conclusion in this case?

Let us recall the quantiles of the standard gaussian distribution $\mathcal{N}(0,1)$ and the η-quantiles $x_{d,\eta}$ of the $\chi^2(d)$ distribution.

If $Z \sim \mathcal{N}(0,1)$, $\begin{cases} \mathbb{P}(Z > 1.645) = 0.95 \\ \mathbb{P}(Z > 1.96) = 0.975 \end{cases}$ and

$x_{d,\eta}$	$\eta = 0.05$	$\eta = 0.95$
$d = 50$	34.764	67.505
$d = 100$	77.929	124.342
$d = 200$	168.279	233.994

Worksheet 4 - Inhomogeneous Poisson process

Exercise 1. Customer's arrival time at a post office.

We count the number of customers who arrive at a post office that opens at 8:00 a.m. and closes at 5:00 p.m.

- From 8:00 a.m. to 11:00 a.m., customers seem to arrive, on average, at a steadily increasing rate that starts with an initial rate of 5 customers per hour at 8:00 a.m., and reaches a maximum of 20 customers per hour at 11:00 a.m.
- From 11:00 a.m. to 1:00 p.m., the (average) rate seems to remain constant at 20 customers per hour.
- Then, the (average) arrival rate drops steadily from 1:00 p.m. until closing time at which time, it has the value of 12 customers per hour.

1. What is a good probability model? Write the intensity.
2. What is the probability that no customer arrive between 8:30 a.m. and 9:30 a.m.?
3. What is the expected number of arrivals between 8:30 a.m. and 9:30 a.m.?

Exercise 2. Repairable System (from the 2021/2022 Exam).

We study a repairable system for which,

- at each failure, the system is repaired to its condition just before failure,
- the repair times can be neglected.

We model the failure times by an inhomogeneous Poisson process with the following intensity:

$$\forall t \geq 0, \qquad \lambda(t) = \begin{cases} 10 - 4t & \text{if } 0 \leq t < 2, \\ 2 & \text{if } 2 \leq t < 12, \\ t - 10 & \text{if } t \geq 12. \end{cases}$$

Denote N_t the number of failures by time $t > 0$ (in months).

1. Recall the definition of an inhomogeneous Poisson process with intensity function $\lambda(\cdot)$.
2. Represent the intensity function and comment its shape.
3. Compute the cumulative intensity function $\Lambda(\cdot)$ of the Poisson process.
   (Indication: you may distinguish 3 cases and check that $\Lambda(2) = 12$ and $\Lambda(12) = 32$).
4. Compute the probability that at least one failure occurs during the first month.
5. Compute the expected number of failures between 2 and 12 months.
6. Compute the probability that at least two failures occur by the third month knowing that there was exactly one failure during the first two months.

Exam (Friday, October 14, 2022)

Duration: 1 hour 30 min

No documents nor calculator are allowed. The parts are independent. It is encouraged (but not mandatory) to provide answers in English. The presentation and the justifications will be taken into account in the assessment. We consider the notation of the Lecture notes.

Exercise 1. Road traffic modeling.

In 2015, a local journal announced "*Le péage autoroutier de Montauban-Sud : l'une des plus belles usines à bouchons de France*". As a response, a study is conducted by the local highway company in order to optimize the traffic fluidity at the toll station[1] of *Montauban-Sud*.

Denote N_t the number of vehicles that arrived to the toll station by time t (in hours, from the beginning of the study). We assume in this exercise that the counting process $(N_t)_{t\geq 0}$ can be modeled by a Poisson process.

Part A - Homogeneous Poisson process.

In this first part, we may assume that the Poisson process is homogeneous, with rate λ.

1. Why is this hypothesis arguable?
2. Define an algorithme for simulating a homogeneous Poisson process with rate λ on $[0, T]$.
3. According to the highway company, the number of tollgates[2] is insufficient if the average number of arrivals during 1 minute exceed 45. They decide to test at level $\alpha = 1\%$

$$\mathcal{H}_0 : \lambda \leq \lambda_0 \quad \text{v.s.} \quad \mathcal{H}_1 : \lambda > \lambda_0.$$

   (a) Determine the value of λ_0.
   (b) What risk is controlled here?
      i. the risk of building an unnecessary tollgate?
      ii. the risk of not building a necessary tollgate?
   (c) We observe the number of arrivals during four working weeks (that is $24 \times 5 \times 4 = 480$ hours). Based on this count, construct a test of $\mathcal{H}_0$ against $\mathcal{H}_1$ with asymptotic size α. Please make clear the estimator, the pivotal statistic, the test statistic and the critical region.
   (d) Define the p-value and express it w.r.t. the cumulative distribution function (c.d.f.) of a distribution to be determined. We obtain a p-value equal to 3.75×10^{-5}. Should the highway company build a new tollgate?

4. Assume here that the intensity is known equal to $\lambda = 3000$. We now classify each vehicle according to its category:
   - type I: cars and motorbikes (in 90% of the cases),
   - type II: trucks (in 10% of the cases).

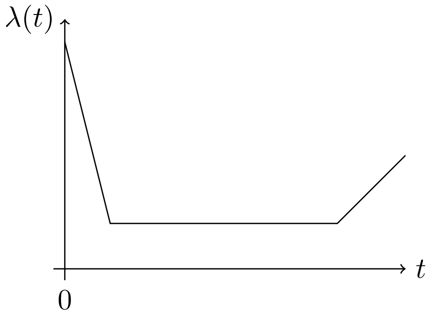
   Denote N_t^{I} (respectively N_t^{II}) the number of cars and motorbikes (respectively of trucks) that arrived to the toll station by time t (in min).

   (a) What is the expected number of trucks per hour?
   (b) What is the probability that at least 2 trucks arrived during 6 min?
   (c) There has been 30 arrivals during 6 minutes. What is the probability that 3 of them where trucks? Prove that this probability does not depend on λ.

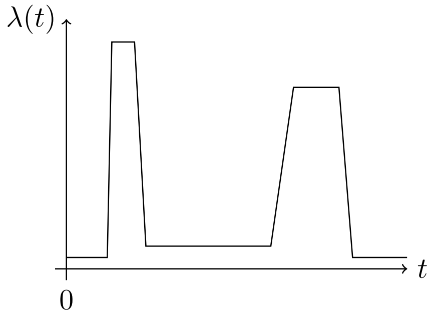
[1] toll station: *gare de péage* in French
[2] tollgate: *poste de péage* in French

Part B - Inhomogeneous Poisson process. In this second part, we may assume that the Poisson process is inhomogeneous, with intensity function $\lambda : (0, +\infty) \to (0, +\infty)$.

1. According to you, which form of intensity should we choose to model the inhomogeneous Poisson process N_t during a day? Justify.



Answer A



Answer B

2. Define the cumulative intensity function.
3. Express the expected number of vehicles that arrived between 8:00 a.m. and 8:30 a.m. as a function of the cumulative intensity.
4. Express the probability that at least 150 vehicles arrive between 6:00 and 7:00 a.m. knowing there where 10 vehicles between 6:00 a.m. and 6:06 a.m. as a function of the cumulative intensity.

Introduction to Poisson processes with R

Objectives

The aim of this session is to manipulate and illustrate the notions introduced in the lecture on Poisson processes with the R software.

Load the `R` file **TP-PoissonProcess-etud.R** (to be completed), available on the Moodle page, and fill in the gaps during this session.

1 Homogeneous Poisson processes observed on a fixed window

First, we consider the case of a fixed observation window (and thus a random number of points).

1.1 Simulation

i. Recall the conditional distribution of the arrival times of a homogeneous Poisson process with rate `lambda` on an interval `[0,Tmax]`, given the number of points in that interval.

ii. Write an `R` function `simulPPh1` with arguments `lambda` and `Tmax`, which simulates such a process and returns the corresponding arrival times.

iii. For `lambda=2` and `Tmax=10`, simulate a homogeneous Poisson process and plot both the counting process and the arrival times.
Indication : you may use the `R` functions `plot()` (with the option `type="s"`), `points()` and `lines()`.

1.2 Maximum likelihood estimator

i. Write an `R` function `MLE1` which returns the Maximum Likelihood Estimator of a homogeneous Poisson process `PPh` observed on a fixed window [0,`Tmax`].

ii. Apply it on different simulated data. What do you observe ?

1.3 Asymptotic behavior of the MLE

In this section, we illustrate the asymptotic behavior [1], as T goes to $+\infty$, of the MLE $\hat{\lambda}_T$ of the rate λ of a homogeneous Poisson process on $[0, T]$.

1.3.1 Strong LLN-type result

First let us illustrate the almost-sure convergence :

$$\hat{\lambda}_T \xrightarrow[T\to+\infty]{a.s.} \lambda. \tag{1}$$

Note that in the lecture, we considered a weak LLN-type result since we only proved the convergence in probability. To illustrate (1) :

i. Fix a rate `lambda=2` and a sequence of `Tmax` that tends to $+\infty$, say `Tillustr=1:500`.

ii. For each `Tmax` in `Tillustr`, simulate a homogeneous Poisson process with rate `lambda` on [0,`Tmax`], and compute the MLE.

iii. Plot the obtained values for the MLE as a function of `Tmax`.

What do you observe ?

1. LLN refers to "Law of Large Numbers" and CLT stands for "Central Limit Theorem".

1.3.2 CLT-type result

We now illustrate the following result :

$$\sqrt{T}(\hat{\lambda}_T - \lambda) \xrightarrow[T\to+\infty]{\mathcal{L}} \mathcal{N}(0,\lambda). \tag{2}$$

To do so, fix `lambda=2` and do the following for different values of `Tmax`.

i. Fix the number of simulations `K=1000` and create a vector of size `K` : `Z=rep(0,K)`.

ii. Store in `Z` a `K`-sample with same distribution as $\sqrt{T}(\hat{\lambda}_T - \lambda)$:

```
for(k in 1:K)
{
  pph=simulPPh1(lambda,Tmax)
  mle=MLE1(pph,Tmax)
  Z[k]=sqrt(Tmax)*(mle-lambda)
}
```

iii. **With the density function :** Plot the histogram of the sample `Z` (which approximates the density of the Z_k's), and compare it with the density of the limit distribution $\mathcal{N}(0,\lambda)$:

```
hist(Z,freq=FALSE,main=paste("Tmax",Tmax,sep="="))
curve(dnorm(x,mean=0,sd=sqrt(lambda)),
      col="red",add=TRUE)
```

What do you observe when `Tmax` equals 1, 10, 100 and 500 ?

iv. **With the c.d.f. :** Plot the empirical cumulative distribution function of `Z`, and compare it with the cumulative distribution function of a $\mathcal{N}(0,\lambda)$:

```
plot(ecdf(Z),main=paste("Tmax",Tmax,sep="="))
curve(pnorm(x,mean=0,sd=sqrt(lambda)),
      col="red",lwd=2,add=TRUE)
```

What do you observe when `Tmax` equals 1, 10, 100 and 500 ?

1.4 Statistical inference : hypothesis testing

Consider a given rate $\lambda_0 > 0$ to test. Given the observation of a homogeneous Poisson process with (unknown "true") rate λ observed on a fixed window `[0,T]`, we aim at testing

$$\mathcal{H}_0 : \lambda = \lambda_0 \quad \text{against} \quad \mathcal{H}_1 : \lambda \neq \lambda_0.$$

i. Construct a test of $\mathcal{H}_0$ against $\mathcal{H}_1$, and express the corresponding p-value in terms of the standard gaussian c.d.f.

ii. Write an `R` function `test1`, with arguments the observed homogeneous Poisson process `PPh`, the observation time `Tmax` and the rate `lambda0` to test, and which returns the p-value of the test constructed in the previous question.

iii. To validate this test on simulated data, we estimate for different values of λ, the probability (or proportion)

$$p(\lambda) = \mathbb{P}_\lambda(\text{reject } \mathcal{H}_0).$$

In particular, if λ satisfies $\mathcal{H}_0$ (i.e. $\lambda = \lambda_0$), then $p(\lambda)$ is the size of the test, and otherwise, it is the power of the test against the alternative λ.

The `R` function `plot.level.power1` in the file **TP-PoissonProcess-etud.R** plots for different values of the rate λ (in `TrueLambda`) confidence intervals for the proportion $p(\lambda)$.

Now, fix `alpha=0.05`, `nsimu=1000` and let `lambda0=2` and `TrueLambda=c(1,1.5,1.8,1.9,2,2.1,2.2,2.5,3)`.
Apply this function for `Tmax = 1,10,100` and `500`. Understand and comment the obtained graphs.

2 Homogeneous Poisson processes with fixed number of points

Second, we consider the case of a fixed number of points (and thus a random observation window).

2.1 Simulation

i. Recall the distribution of the interarrival times of a homogeneous Poisson process with rate `lambda`.

ii. Write an `R` function `simulPPh2` with arguments the rate `lambda` and the number of points `n`, which simulates such a process and returns the corresponding arrival times.

iii. As in the "fixed window" case, fix `lambda=2` and `n=20`, simulate a homogeneous Poisson process and plot both the counting process and the arrival times.

2.2 Maximum likelihood estimator

i. Write an `R` function `MLE2` which returns the maximum likelihood estimator of the rate of a homogeneous Poisson process `PPh` observed up to the `n`th point.

ii. Apply it on different simulated data. What do you observe ?

2.3 Asymptotic behavior of the MLE

As in section 1.3, illustrate the asymptotic behavior of the MLE $\hat{\lambda}_n$ of a homogeneous Poisson process with rate λ observed up to the nth point as n goes to $+\infty$, that is

i. (Strong) LLN-type result :

$$\hat{\lambda}_n \xrightarrow[n\to+\infty]{a.s.} \lambda.$$

ii. CLT-type result :

$$\sqrt{n}(\hat{\lambda}_n - \lambda) \xrightarrow[n\to+\infty]{\mathcal{L}} \mathcal{N}(0, \lambda^2).$$

2.4 Statistical inference : confidence intervals

i. Recall both asymptotic and non-asymptotic confidence intervals for the (unknown) rate `lambda` of a homogeneous Poisson process `PPh` observed up to the `n`th point.

ii. Write an `R` function with arguments `PPh`, `alpha=0.05` and a boolean `asymptotic`, and which, depending on the value of `asymptotic` (`TRUE` or `FALSE`), returns the corresponding confidence interval for `lambda`.

iii. As for the testing problem, we want to validate the confidence intervals on simulated data.
To do so, illustrate, in both asymptotic and non-asymptotic cases, the fact that the proportion of times the rate belongs to the obtained confidence interval is larger than the fixed confidence level `1-alpha`.
You may fix `lambda=2`, and consider `n=10` or `n=100`.

3 Inhomogeneous Poisson processes

We now aim at simulating inhomogeneous Poisson processes with given intensity function, on a given interval.

i. Recall the thinning algorithm. In particular, when does it apply ?

ii. Write an `R` function `simulPPi` with arguments the intensity function `lambda_fct`, `Tmax` and an upper bound `M`, which simulates an inhomogeneous Poisson processes with intensity function `lambda_fct()` on `[0,Tmax]`, and returns the corresponding arrival times.

iii. Simulate and represent inhomogeneous Poisson processes on $[0, 10]$ with intensity functions :
- $\lambda_1 : t \mapsto 2 \times \mathbb{1}_{[0,7]}(t) + 8 \times \mathbb{1}_{]8,10]}(t)$.
- $\lambda_2 : t \mapsto 2t$.
- $\lambda_3 : t \mapsto \ldots$.

Bibliography

[CT97] Christiane Cocozza-Thivent. *Processus stochastiques et fiabilité des systèmes*, volume 28. Springer Science & Business Media, 1997.

[FF04] Dominique Foa and Aimé Fuchs. *Processus stochastiques, processus de Poisson, chaînes de Markov et martingales.* Dunod, Paris, 2004.

[Mik09] Thomas Mikosch. *Non-Life Insurance Mathematics: An Introduction with the Poisson Process.* Springer-Verlag Berlin Heidelberg, 2nd edition, 2009.

[Ros06] Sheldon M Ross. *Introduction to Probability Models.* Academic press, 9th edition, 2006.