

Poisson processes and application to reliability theory and actuarial science

Academic year 2021-2022

Mélanide ALBERT

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

Bibliography

- [CT97] Christiane Coccozza-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.

Contents

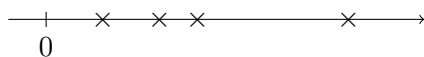
1. Some reminders in probability	7
I. Poisson random variable	7
II. Exponential random variable	8
III. Memoryless distribution	11
IV. Hazard rate	13
V. On the gamma and Erlang's distributions	15
2. Homogeneous Poisson processes	19
I. Definitions	19
I.1. Point processes and counting processes	19
I.2. Homogeneous Poisson process on \mathbb{R}_+	23
II. Arrival and interarrival times	27
II.1. First arrival time after a given instant	27
II.2. Exponential interarrival times	27
II.3. Conditional distribution of the arrival times	30
II.4. Construction of a homogeneous Poisson process	33
III. Divisibility of homogeneous Poisson processes	35
III.1. Two states divisibility	35
III.2. Infinite divisibility	37
3. Statistics for homogeneous Poisson processes	39
I. Introduction	39
II. Fixed number of points, random observation time	39
II.1. Maximum Likelihood Estimator (MLE)	40
II.2. Non-asymptotic properties of the MLE	41
II.3. Asymptotic distribution of the MLE	42
II.4. Confidence intervals and statistical tests	43
III. Fixed window, random number of points	45
III.1. Maximum Likelihood Estimator	45
III.2. Non-asymptotic properties of the MLE	46
III.3. Asymptotic distribution of the MLE	48
III.4. Confidence intervals and statistical tests	50
4. Inhomogeneous Poisson processes	53
I. Definitions	53
II. Construction of an inhomogeneous Poisson process	56
II.1. Time change	56

II.2.	Acceptance/Rejection (or Thinning)	61
III.	Arrival times	63
III.1.	First arrival time after a given instant	63
III.2.	(Conditional) distribution of the arrival times	63
IV.	Statistics for inhomogeneous Poisson processes	66

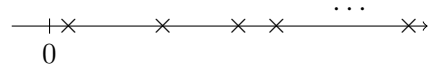
Introduction

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 from 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 occurrence 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?
 - ...

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 occurrence. 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:

$$\forall t \leq 0, \quad \mathbb{E}[e^{tN}] = \exp(\lambda(e^t - 1)).$$

2. The Poisson distribution is also characterized by its characteristic function:

$$\forall t \in \mathbb{R}, \quad \mathbb{E}[e^{itN}] = \exp(\lambda(e^{it} - 1)).$$

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

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

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 $\mathbb{E}[N] = \text{Var}(N) = \lambda$.

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

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

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

□

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 $t < \lambda$,

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

2. Its distribution functions equals: $F : t \in \mathbb{R} \mapsto (1 - e^{-\lambda t}) \mathbb{1}_{\{t \geq 0\}}$.

3. Its survival function satisfies: $\forall t \geq 0, \mathbb{P}(X > t) = e^{-\lambda t}$.

Proof. In exercise.

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

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 $t \geq 0$.

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 $\text{Var}(X) = 1/\lambda^2$.

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

□

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:* $\mathbb{P}(X_1 < X_2) = \mathbb{E}[\mathbf{1}_{\{X_1 < X_2\}}]$.

□

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 $\min_{1 \leq i \leq n} X_i$.

□

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 $t, s > 0$,

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

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 it 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 $\mathbb{P}(X > t + s | X > t)$.
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.
 - b) Let $s, t > 0$. Prove that $G(t + s) = G(s)G(t)$.
 - c) Deduce from Lemma 1.11, that for all $t > 0$, $G(t) = e^{-\lambda t}$ where λ is a well-chosen constant (to determine).
 - d) Conclude.

□

Comment 1.12. Discrete memoryless distributions are geometric distributions.

IV Hazard rate

In this section, we only consider non-negative continuous real-valued random variables.

Definition 1.13.

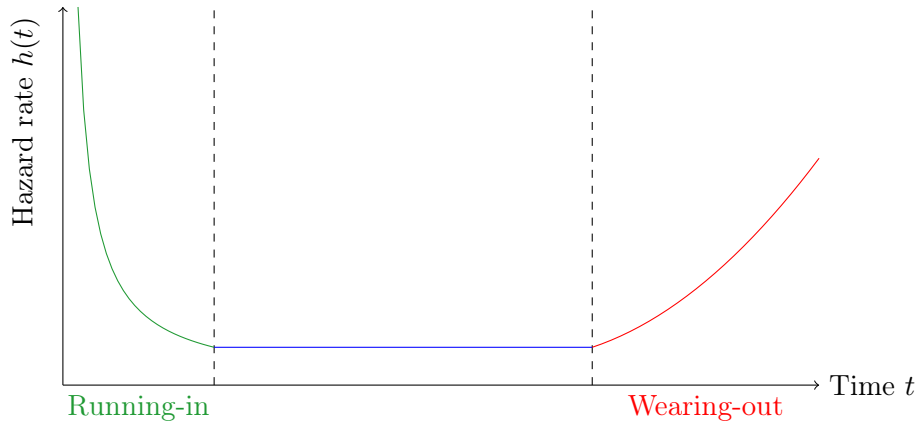
Let X be a non-negative real random variable with density function f w.r.t. the Lebesgue measure. The hazard rate function h is defined by

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

Example: 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 time to failure of a component, the hazard rate corresponds to the probability of a failure occurring just 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 probability of failure decreases with the age. On the contrary, if the failure rate h is increasing, it means that the probability of failure increases with the age. In the reliability theory, the failure rate is shaped as below.



Proposition 1.14.

Let X is a continuous r.v. Then 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 .

\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

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

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.

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

\Leftarrow Assume that there exists a function h with 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 $f : t \mapsto f(t) = h(t)G(t)$.

ii) Deduce the hazard rate of X .

□

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.

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

⇐ 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.

□

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 $t < \lambda$,

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

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

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

Proof. In exercise.

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

- 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

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

□

