UNIVERSITAT DE BARCELONA
Facultat de Matemàtiques i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

The mathematical structure of the Poisson process

Autora: Clara Ribas Torres

Director: Dr. Francesc Xavier Massaneda Clares
Realitzat a: Departament de Matemàtiques i Informàtica

Barcelona, 10 de juny de 2025

Abstract

This thesis explores Poisson processes as fundamental models for random, independent events occurring over time or space. After introducing their core properties, including the Poisson distribution and spatial independence, we develop key theoretical results: the Superposition, Mapping, and Existence Theorems. The work then presents essential analytical tools like Campbell's and Rényi's theorems for studying these processes. Together, these provide a complete framework for modeling systems where events occur unpredictably but with consistent average rates.

Resum

Aquesta tesi explora els processos de Poisson com a models fonamentals per a esdeveniments aleatoris i independents que tenen lloc al llarg del temps o en l'espai. Després d'introduir-ne les propietats bàsiques, incloent-hi la distribució de Poisson i la independència espacial, es desenvolupen resultats teòrics clau: els teoremes de Superposició, de Mapeig i d'Existència. A continuació, es presenten eines analítiques essencials com els teoremes de Campbell i de Rényi per a l'estudi d'aquests processos. En conjunt, aquests elements ofereixen un marc complet per modelar sistemes en què els esdeveniments es produeixen de manera imprevisible però amb una taxa mitjana constant.

2020 Mathematics Subject Classification. 60G55, 60G50, 28A33

Agraïments

Abans de començar m'agradaria agrair a totes aquelles persones que m'han animat i donat suport durant aquests mesos perquè aquest treball veiés finalment la llum.

Vull agrair-me, en primer lloc, a mi mateixa per haver afrontat aquest repte. Escriure aquest treball lluny de casa, amb l'enyorança a sobre i en un entorn nou, no ha estat fàcil. Hi ha hagut moments de dubte, hores de dedicació en lloc d'aprofitar els últims mesos a Amsterdam, i dies en què la feina semblava no avançar. M'alegro d'haver-me escoltat quan vaig decidir no rendir-me, d'haver après a treballar fins i tot quan la motivació minvava, i d'haver donat el millor perquè aquest projecte fos una realitat.

No hagués estat possible trobar aquesta força en mi sense el suport incondicional de la meva família i de la meva parella. Gràcies per creure en mi fins i tot quan jo mateixa dubtava, per recordar-me sempre que amb esforç i il · lusió, no hi ha res que no pugui aconseguir.
I especialment a tu, Iván: gràcies per acompanyar-me pas a pas fins a la meta malgrat la distància. Gràcies per saber veure tot aquell potencial que porto dins i que ni jo mateixa sempre aconsegueixo veure. Per fer-me sentir que Amsterdam i Sarajevo no són tan lluny.

Gràcies al Dr. Francesc Xavier Massaneda Clares per la seva ajuda i guia en aquest darrer tram de carrera. Des del primer moment em va animar a emprendre aquesta aventura i, tot i que el procés ha estat a través d'una pantalla, sempre ha estat disponible per a tot el que pogués necessitar. Gràcies per respondre amb paciència les meves preguntes, per petites que fossin, i per explicar-me amb tant de detall i ilusió tot allò que no entenia. Tant de bó algun dia pogués saber la meitat del que tu saps. És admirable.

Finalment, vull donar les gràcies al meu avi Miquel. Tot i que no pugui veure aquet treball, sé que allà on sigui, se sent orgullós del que he fet. Algún dia seré tan intel · ligent com tu em feies sentir. Ara que he entregat i presentat el treball, ja puc contestar a la teva gran pregunta: "Com van es negocis per Barcelona?" I jo et dic el que tant desitjaves sentir: "Finalment, tens una néta matemàtica".

Contents

Introduction

Completely independent random sequences, a fundamental construction in probability theory, model a broad range of real-world phenomena, with the Poisson process standing as one of the most prominent and widely utilized examples. In general, in statistics and probability theory, a random collection of points in a given space S is known as a *point process.*

Informally, the Poisson process is a particular type of process consisting of discrete, independent events distributed over a continuous domain. This document explores the formal definition of the Poisson process, its foundational properties, and key characterizing features.

Specifically, the *Poisson process* defined over a space S is a random sequence of points Π satisfying two fundamental properties. For a set $A \subseteq S$, let $N(A)$ denote the random variable counting the number of points falling in A, that is $N(A) = \#(\Pi \cap A)$. The defining properties are:

a) **Independence:** The number of points falling in disjoint regions are independent random variables. That is, if A and B are disjoint subsets of the space S, then the random variables $N(A)$ and $N(B)$ are independent.

b) **The Poisson distribution:** Given a subset $A \subseteq S$, the variable $N(A)$ follows a Poisson distribution. Therefore, there exists a non-negative parameter $\mu(A)$ such that

$$\mathbb{P}(N(A) = k) = e^{-\mu(A)} \cdot \frac{(\mu(A))^k}{k!}, \quad k = 0, 1, 2, ...$$

By the properties of the Poisson distribution, the parameter $\mu(A)$ coincides with the average number of points falling in A.
As we shall see, at least in regular situations, independence in disjoint regions forces the counts to be Poisson distributed, because it's the only distribution

that preserves independence while modeling discrete events. This means that condition b) is natural if we have condition a).

As it turns out, the values $\mu(A)$ gives rise to a measure in the underlying space S, called the *intensity of the Poisson process*. Conversely, and more importantly, for any σ-finite, non-atomic measure μ on S, there exists a Poisson process with intensity μ. The intensity describes how points are distributed across space. It gives the expected number of points in any given region, reflecting how frequent they are. Higher intensity means more points are likely to occur in that area.

The *state space* S mentioned above is usually a Euclidean space or a manifold of some dimension d. It must be a measurable space with a σ-field of measurable sets, ensuring that individual points are distinguishable. When $S = \mathbb{R}^d$, the measurable sets are always taken to be the Borel sets, those belonging to the smallest σ-field containing open sets.

When $S = \mathbb{R}^d$, the mean measure is in most interesting cases given in terms of a *rate* or *intensity function*. This is a positive function λ on S that describes the density of points in space. In this case, μ is given by integrating λ with respect to the d-dimensional Lebesgue measure:

$$\mu(A) = \int_A \lambda(x)dx.$$

If $\lambda(x)$ is constant, then μ is just a multiple of the Lebesgue measure, and the Poisson process is known as *uniform* or *homogeneous*. This random process has stochastic properties which are unchanged under translations or rotations of S, and is the only Poisson process with this property. In any other case, the Poisson process is called *non-homogeneous* or *non-uniform*. Then, the probability does depend on the location of the particular region and, in general, it will not be the same if the set is rotated or translated.

The Poisson distribution was first developed by Siméon Denis Poisson in 1830 to predict the number of wins a gambler might expect over many games. Though initially applied to gambling, its broader usefulness became clear during World War II, when British statistician R.D. Clarke used it to show that German bombings in London were distributed as a homogeneous Poisson process. Since then, the Poisson distribution has found applications in many fields, including astronomy, biology, telecommunications, ecology, physics, and more. [3] [2]

On the real line, the Poisson process models random events occurring over time. One of its most prominent applications is in queuing theory [1], where

it is widely used to model customer arrivals in service systems such as banks or call centers. In this context, arrivals are assumed to occur randomly and independently at a constant average rate μ. The number of arrivals in a given time interval follows a Poisson distribution, while the time between arrivals is exponentially distributed. This framework provides powerful tools for analyzing key performance metrics, including waiting times, queue lengths, and service efficiency. In two-dimensional space, the Poisson point process, also known as the spatial Poisson process, extends this model to represent the random locations of scattered objects, such as particles hitting a detector, or trees distributed throughout a forest. [9]

Despite its power and simplicity, the Poisson process has limitations: it does not account for interactions between points, such as attraction or repulsion. This can lead to its overuse in scenarios where such dependencies matter. For instance, modeling WiFi router placement with a Poisson process ignores the deliberate spacing to prevent interference and ensure coverage. In such cases, more advanced point processes are needed to accurately reflect spatial interactions.

Below is the organizational structure of the memoir. The opening chapter lays the mathematical groundwork for Poisson processes by introducing measurable spaces, σ-algebras, and non-atomic measures.
It then presents the process's additive property: given $A, B \subseteq S$ with no intersection, then $N(A \cup B)$ follows a Poisson distribution of parameter

$$\mu(A \cup B) = \mu(A) + \mu(B).$$

This extends naturally to the temporal setting: if time is divided into non-overlapping intervals, the number of events in each interval is independent of the others, and the total number of events over the combined interval is simply the sum of the events in each part.
Additionally, the Poisson process is closely connected to the binomial distribution through a limiting relationship.
Finally, we present a heuristic argument demonstrating that the assumption of independence naturally leads to the conclusion that the variables $N(A)$ must follow a Poisson distribution. This fundamental property is central to the Poisson process and explains its broad applicability in modeling random phenomena across both space and time.

Chapter 2 focuses on several natural results that highlight the versatility of the Poisson process. First, the superposition theorem (Theorem 2.2) shows that merging independent Poisson processes creates another Poisson

process and the resulting intensity is just the sum of the individual intensities. The superposition theorem is natural for the Poisson process because of its key properties. When multiple independent Poisson processes are combined, their randomness (independence) remains intact. Also, the sum of independent Poisson variables is a Poisson variable. This simple additive behavior does not hold for other point processes, which often involve dependencies, interactions, or spatial constraints that are disrupted by superposition, making the resulting process no longer follow the same structure.

Next, the mapping theorem (Theorem 2.4) states that applying a measurable transformation, with mild restrictions, to a Poisson process results in another Poisson process, in general with a different intensity. This new intensity is given by the push-forward of the original measure, which reflects how the intensity is redistributed under the transformation. Despite the change in space or coordinates, such as projection or rescaling, the points remain independently and randomly distributed. This property demonstrates how reliable Poisson processes are for modeling random events in various situations.

In the final part of Chapter 2, we prove the Existence Theorem (Theorem 2.8).The key insight is that for any subset $B \subseteq S$, the Poisson process restricted to B and conditioned on $N(B) = n$ follows a multinomial distribution with parameters n and $p(A) = \frac{\mu(A)}{\mu(B)}$ for subsets $A \subseteq B$. This property paves the way to proving one of this thesis's central results. This theorem guarantees that for any σ-finite and non-atomic measure μ on S, there exists a Poisson process with intensity μ. The formal construction of this result is presented in Section 2.4. The existence theorem is powerful because it guarantees that Poisson processes can be built with nearly any reasonable measure μ serving as their average spatial distribution.

The last chapter develops the fundamental analytical tools for studying Poisson processes through the concept of the so-called linear statistics, which are random sums of the form

$$\Lambda_\Pi(x) = \sum_{x \in \Pi} f(x),$$

where f is a measurable function. The goal is to understand both the convergence conditions for these sums and their statistical properties, such as expectation, variance, and full distribution.
The core theoretical framework emerges through Campbell's Theorem (Theorem 3.1), which establishes the relationship between the expectation of a

function summed over a point process and an integral involving the mean measure of that process. Specifically, for a point process Π defined on a space S with mean measure μ and a linear statistic Λ_Π as above, its moment generating function is given by

$$\mathbb{E}(e^{\theta\Lambda}) = \exp\left(\int_S (e^{\theta\Lambda(x)} - 1)\, d\mu(x)\right).$$

This formula is significant because it enables the calculation of expected values and variances of random sums indexed by the point process, characterizes the distribution of the process, and provides conditions under which such sums converge almost surely.

The chapter finishes by revealing a deep characterization theorem. Rényi's theorem states that the avoidance probabilities

$$\mathbb{P}(N(A) = 0) = e^{-\mu(A)} \quad A \subseteq S$$

alone suffice to determine a Poisson process. This connects elegantly to Kendall's foundational work on random sets and highlights the duality between two perspectives: modeling via counting measures versus characterizing through pointwise avoidance functions.

Chapter 1

Foundations and properties of the Poisson process

In this chapter, we first introduce the foundational concepts required to rigorously define the Poisson process. These include measurable and probability spaces, non-atomic and σ-finite measures, and random variables.

Following this, we will discuss the additive property of the Poisson process and its connection to the binomial distribution via limiting behavior. Finally, we will highlight the importance of the independence property, which plays a key role in the Poisson process and its applications.

1.1 Probability structures for Poisson processes

In general, we will consider Poisson processes in a measurable space S (usually $\mathbb{R}^d$, or some subset of $\mathbb{R}^d$). Together with S we have a σ-algebra Σ of subsets of S, and a non-negative measure $\mu : \Sigma \to [0, +\infty)$. These form a measure space (S, Σ, μ).

Recall that a σ-algebra Σ is a family of subsets such that:

i) $S \in \Sigma$.

ii) Σ is closed under complementary: if $A \in \Sigma$, then its complementary $A^c = \Omega \setminus A$ is also in Σ.

iii) Σ is closed under countable unions: if $A_1, A_2, A_3, \ldots$ are in Σ, then their union $\cup_{i=1}^{\infty} A_i$ is also in Σ.

On the other hand, a measure is a function $\mu : \Sigma \to [0, +\infty)$ such that:

i) $\mu(\emptyset) = 0$

ii) σ-additivity: if $\{A_n\}_{n=1}^{\infty}$ is a countable collection of pairwise disjoint subsets of Σ, then

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

Definition 1.1. *A measurable space is a triple (S, Σ, μ) where S is a non-empty set, Σ is a σ-algebra of subsets of S and $\mu : \Sigma \to [0, +\infty)$ is a measure. The members of Σ are called μ-measurable sets.*

Example 1.2. The canonical example is given by $S = \mathbb{R}^d$, Σ the Lebesgue measurable sets, and $\mu = |\cdot|$ the Lebesgue measure.

Example 1.3. *Didac's measure on a point:* Fix $x_0 \in \mathbb{R}^d$. Let $\Sigma = \mathcal{P}(\mathbb{R}^d)$ and define the measure

$$\delta_{x_0} : \mathcal{P}(\mathbb{R}^d) \to [0, +\infty)$$

as follows: for $A \in \mathcal{P}(\mathbb{R}^d)$

$$\delta_{x_0}(A) = \begin{cases} 1, & \text{si } x_0 \in A, \\ 0, & \text{si } x_0 \notin A. \end{cases}$$

Let us check that δ_{x_0} satisfies the two conditions in the definition of measure. Clearly it satisfies property i), since by definition $\delta_{x_0}(\emptyset) = 0$. Therefore, we only have to prove condition ii). If $x_0 \notin \cup_j E_j$ then

$$\delta_{x_0}(\cup_j E_j) = \sum_j \delta_{x_0}(E_j) = \sum_j 0 = 0.$$

On the other hand, if $x_0 \in \cup_j E_j$ then there exists a unique j_0 such that $x_0 \in E_{j_0}$. Thus,

$$\delta_{x_0}(\cup_j E_j) = \delta_{x_0}(E_{j_0}) + \sum_{j \neq j_0} \delta_{x_0}(E_j) = 1 + 0 = 1.$$

We will need to avoid cases where the probability that the Poisson process contains a given point is positive. This is achieved by considering underlying measures for which the mass of each single point is zero.

Definition 1.4. *A measure μ is called non-atomic if $\mu(\{x\}) = 0$, for any $x \in S$.*

The Lebesgue measure is the most well-known example of a non-atomic measure. Any measure of the form $\mu(A) = \int_A \lambda(x)dx$, as given in the introduction, is non-atomic. From now on, in the context of the Poisson process, we assume that the measure μ is non-atomic.

Additionally, we require μ to be σ-finite, ensuring that the associated Poisson process does not have accumulation points in any region of S.

Definition 1.5. *A measure μ defined on a measurable space (S, Σ) is called σ-finite if there exists a countable collection of measurable sets $\{A_n\}_{n=1}^{\infty}$ such that:*

1. $S = \cup_{n=1}^{\infty} A_n$,
2. $\mu(A_n) < \infty$ *for all* $n \in \mathbb{N}$.

A very particular case of measurable space is a *probability space*, a triple $(\Omega, \mathcal{F}, \mu)$, where Ω is a set of elementary outcomes of a random experiment, $\mathcal{F}$ is a σ-algebra of subsets of Ω (called events), and $\mu : \mathcal{F} \to [0, 1]$ is a measure, denoted as $\mu = \mathbb{P}$, such that $\mathbb{P}(\Omega) = 1$.

Now that the measurement conditions have been introduced, we will introduce random variables, and in particular the counting random variables $N(A),\ A \in \Sigma$.

Definition 1.6. *A real-valued random variable is a function $X : \Omega \to \mathbb{R}$ that is measurable, meaning that for every $x \in \mathbb{R}$ the set $\{\omega; X(\omega) \leq x\}$ is included in $\mathcal{F}$.*

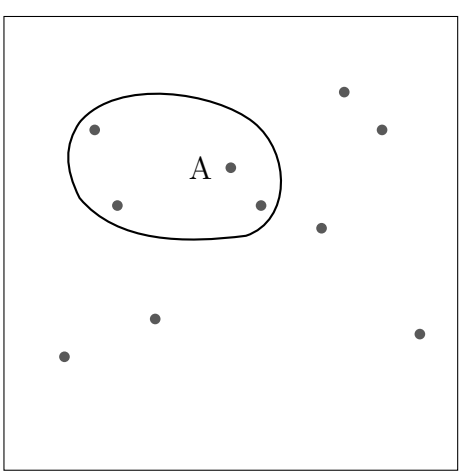
Our goal is to construct a σ-algebra $\mathcal{F}$ such that the counting functions $N(A), A \in \Sigma$ are random variables.

A Poisson process defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and associated with the state measure space (S, Σ, μ) is a mapping Π that assigns to each element of Ω a countable subset of S with certain properties. As a result, the space $S^{\infty} = S \times S \times S \times \cdots$ represents the collection of all possible (countable) sequences in S. A sequence $\{x_k\}_{k=1}^{\infty}$ is considered as an infinite vector in S^{∞}, written as $(x_1, x_2, \ldots, x_n, \ldots)$. The goal is to generate random sequences in

S (point processes) that exhibit specific characteristics. To represent such sequences formally, we denote $\Pi(\omega) = \{x_k(\omega)\}_{k=1}^{\infty}$.

To study these processes, we select $A \in \Sigma$ and consider the function counting the number of points from Π that fall in A

$$N(A) = \#(\Pi(\omega) \cap A) = \#\{k | x_k(\omega) \in A\}.$$



$N(A) = 4$

We want $N(A) : \Omega \to \mathbb{N} \cup \{0\}$ to be a random variable, meaning that we need that the sets

$$E(A, n) := \{\omega \in \Omega : N(A)(\omega) \leq n\}$$

are measurable in Ω, for all $n \in \mathbb{N} \cup \{0\}$. We simply define $\mathcal{F}$ as the σ-algebra generated by the sets $E(A, n)$, with $A \in \Sigma$ and $n = 0, 1, 2, \ldots$.

As explained in the introduction, Π will be a Poisson process if:

1. For any $A, B \subseteq S$ μ-measurable sets with $A \cap B = \emptyset$, the random variables $N(A)$ and $N(B)$ are independent.

2. For any $A \subseteq S$ μ-measurable set, the random variable $N(A)$ follows a Poisson distribution with parameter $\mu(A)$, that is:

$$\mathbb{P}(N(A) = k) = e^{-\mu(A)} \frac{\mu(A)^k}{k!}, \quad k \geq 0.$$

1.2 The Poisson distribution

We now provide a brief reminder of the main properties of the Poisson distribution. These will be useful for later properties of the Poisson process.

A random variable X has the *Poisson distribution* with parameter $\lambda > 0$,

$$X \sim \mathcal{P}(\lambda),$$

if its possible values are non-negative integer numbers and if

$$\mathbb{P}(X = k) := \frac{\lambda^k}{k!} e^{-\lambda}, \quad k > 0. \tag{1.1}$$

It is easy to check that $\mathbb{E}(X) = \mathbb{V}(X) = \lambda$.

It is sometimes useful to extend the definition of $\mathcal{P}(\lambda)$ to include the boundary cases 0 and ∞. So, if $\lambda = 0$, then $\mathbb{P}(X = 0) = 1$, since we take $0^0 = 1$. When $\lambda = \infty$ we put $\mathbb{P}(X = \infty) = 1$.

Everything is known of the Poisson distribution. In particular, the *probability generating function* of X is explicit and given by

$$\mathbb{E}(z^X) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} z^k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{(\lambda z)^k}{k!} = e^{\lambda(z-1)}, \quad z \in [0, 1]. \tag{1.2}$$

1.2.1 Poisson additivity

The most important feature of the Poisson distribution is its additivity: if $X \sim \mathcal{P}(X)$ and $Y \sim \mathcal{P}(Y)$ are independent random variables, then $X + Y \sim \mathcal{P}(X + Y)$. This extends to the sum of any finite number of independent Poisson variables. Let us see this; for integers $r, s \geq 0$

$$\mathbb{P}(X = r, Y = s) = \mathbb{P}(X = r)\mathbb{P}(Y = s) = \frac{\lambda^r e^{-\lambda}}{r!} \frac{\mu^s e^{-\mu}}{s!}.$$

With this, the distribution of the random variable $X + Y$ is given by

$$\begin{aligned} \mathbb{P}(X + Y = n) &= \sum_{r=0}^{n} \mathbb{P}(X = r, Y = n - r) = \sum_{r=0}^{n} \frac{\lambda^r e^{-\lambda}}{r!} \frac{\mu^{n-r} e^{-\mu}}{(n-r)!} \\ &= \frac{e^{-(\lambda+\mu)}}{n!} \sum_{r=0}^{n} \binom{n}{r} \lambda^r \mu^{n-r} = \frac{(\lambda + \mu)^n e^{-(\lambda+\mu)}}{n!}. \end{aligned}$$

Thus, $X + Y$ has distribution $\mathcal{P}(\lambda + \mu)$, as stated.

1.2.2 Poisson as limit of binomials

The *binomial distribution* models a random experiment in which each trial can have only two possible outcomes: success, with probability p, and failure, with probability $1-p$. We consider a sequence of n independent Bernoulli trials, where the probability of success in each trial remains constant. We define the random variable X as the total number of successes obtained in these n trials. The possible values of this variable are the integers from 0 to n. We say that X follows a binomial distribution with parameters n and p, and we denote it as:

$$X \sim \text{Bin}(n, p).$$

The distribution of X is given by

$$\text{Bi}(n, p; k) = \mathbb{P}(X = k) := \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, \ldots, n.$$

It is well-known, and easy to check, that for $X \sim Bin(n, p)$,

$$\mathbb{E}(X) = np \quad \text{and} \quad \mathbb{V}(X) = np(1-p).$$

In the context of the Poisson model, it is useful to study the binomial distribution because, under certain conditions, the binomial distribution can be approximated by the Poisson distribution. Specifically, when the number of trials n is large and the probability of success p is small, such that the product $np = \lambda$ remains constant, the binomial distribution $Bin(n, p)$ converges to a Poisson distribution with parameter λ. This relationship is important because it allows the Poisson distribution to serve as a mathematical simplification of the binomial in situations where we model rare events occurring over a large number of trials. Such scenarios are common in phenomena like product defects, or accidents within a given time interval.

If $\lambda = np$, the average number of successes, is held constant while n becomes large and p small, expression (1.1) appears as a limit: for integer $r \geq 0$

$$\begin{aligned} \lim_{n\to\infty} \text{Bin}\left(n, \frac{\lambda}{n}; r\right) &= \lim_{n\to\infty} \binom{n}{r} (\frac{\lambda}{n})^r (1 - \frac{\lambda}{n})^{n-r} \\ &= \lim_{n\to\infty} \frac{n^r}{r!} \cdot \frac{\lambda^r}{n^r} \cdot e^{-\lambda} = \frac{\lambda^r e^{-\lambda}}{r!}. \end{aligned}$$

1.3 The Strength of Independence

The most determining property of a Poisson process is the independence of the counting functions of disjoint sets. To justify this, let us show heuristically, under mild regularity assumptions, how the property of independence

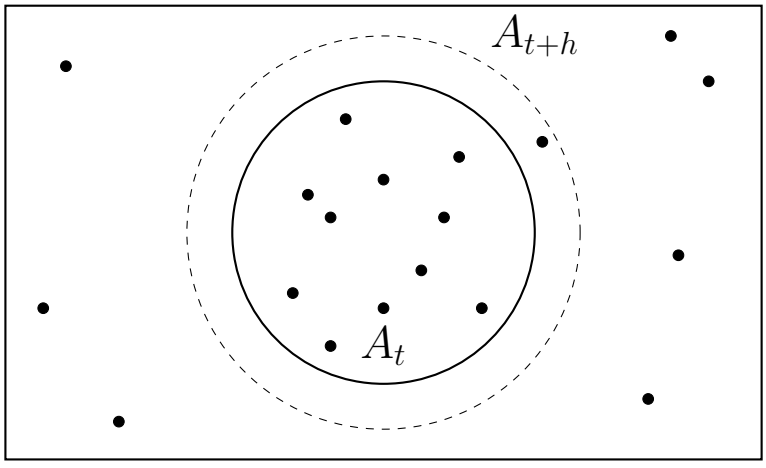
leads to the conclusion that the counting random variables $N(A)$ must follow a Poisson distribution.

Let A_t be a family of test sets parameterized by t, such as a ball of radius $t \geq 0$, a cube of side t, or any increasing family of sets with a regular boundary in $\mathbb{R}^d$. For $n \geq 0$, define the probabilities:

$$p_n(t) = \mathbb{P}(N(A_t) = n), \quad q_n(t) = \mathbb{P}(N(A_t) \leq n) = \sum_{j=0}^{n} p_j(t).$$

Since $N(A_t)$ is a non-decreasing function of t and $q_n(t)$ is non-increasing, both functions exhibit only jump discontinuities, making them differentiable almost everywhere.

The random variable $N(A_t)$ increases from n to $n+1$ whenever the expanding boundary of A_t encounters a new random point. The probability of this happening in the interval $(t, t+h)$ for a small $h > 0$ is approximately the probability that a point lies within the thin region between A_t and A_{t+h}. This probability decreases as h decreases. Also, when expanding from A_t to A_{t+h}, the number of added points will remain finite with probability one, preventing the possibility of capturing an infinite number of them in the process.



We now refine our analysis making a mild assumption: as A_t evolves, new points are captured one by one:

$$\mathbb{E}[N(A_{t+h}) - N(A_t) \geq 2] = o(h) \tag{1.3}$$

as $h \to 0$.

This assumption allows us to estimate the probability of having points in the set difference $A_{t+h} \setminus A_t$.

Claim: Assuming (1.3)

$$\begin{aligned}\mathbb{P}(N(A_{t+h}) - N(A_t) \geq 1) &= \mathbb{E}[N(A_{t+h})] - \mathbb{E}[N(A_t)] + o(h) \\ &= \mathbb{E}[N(A_{t+h} \setminus A_t)] + o(h).\end{aligned}$$

Proof. By definition and the previous assumption,

$$\begin{aligned}\mathbb{E}[N(A_{t+h})] - \mathbb{E}[N(A_t)] &= \mathbb{E}[N(A_{t+h}) - N(A_t)] \\ &= \sum_{k=0}^{\infty} k \cdot \mathbb{P}(N(A_{t+h}) - N(A_t) = k) \\ &= \mathbb{P}(N(A_{t+h}) - N(A_t) = 1) + \sum_{k=2}^{\infty} k \cdot \mathbb{P}(N(A_{t+h}) - N(A_t) = k) \\ &= \mathbb{P}(N(A_{t+h}) - N(A_t) = 1) + \mathbb{E}[N(A_{t+h} - N(A_t) \geq 2].\end{aligned}$$

□

Thus, this last equality reads as:

$$\mathbb{P}(N(A_{t+h}) - N(A_t) = 1) = \mathbb{E}[N(A_{t+h})] - \mathbb{E}[N(A_t)] + o(h)$$

which, defining $\mu(t) = \mathbb{E}[N(A_t)]$, simplifies to:

$$\mathbb{P}(N(A_{t+h}) - N(A_t) = 1) = \mu(t+h) - \mu(t) + o(h).$$

Lemma 1.7. *Under the condition above, the variable $N(A_t)$ follows a Poisson distribution of parameter $\mu(t)$, that is*

$$\mathbb{P}(N(A_t) = n) = p_n(t) = e^{-\mu(t)} \frac{\mu(t)^n}{n!}, \quad n = 0, 1, \ldots.$$

The proof of this lemma is lengthy and technical. It relies on the analysis of a system of differential equations satisfied by the probabilities p and q.

Proof. For a given n, consider the event $\mathrm{E} = \{N(A_t) \leq n\}$. Let's analyze how this event changes when the time progresses from t to $t+h$. Two options can happen:

1. $N(A_{t+h}) \leq n$. Since the probability of having two or more points in $A_{t+h} \setminus A_t$ is assumed to be $o(h)$, this occurs primarily in two cases:

(a) $N(A_t) < n$,

(b) $N(A_t) = n$ (and that means $N(A_{t+h})$ is also n). Then, there are no points in $A_{t+h} \setminus A_t$.

2. $N(A_t) = n$ and there is also a point in $A_{t+h} \setminus A_t$. Thus, $N(A_{t+h}) = n+1$.

Hence, if

$$E_1 = \{N(A_{t+h}) \leq n\} \text{ and } E_2 = \{N(A_t) = n\} \cap \{N(A_{t+h} \setminus A_t) = 1\},$$

then $E = E_1 \cup E_2$ with $E_1 \cap E_2 = \emptyset$. Therefore,

$$\mathbb{P}(E) = \mathbb{P}(E_1) + \mathbb{P}(E_2).$$

Since A_t and $A_{t+h} \setminus A_t$ are disjoint, the assumption of independence implies that $N(A_t)$ and $N(A_{t+h} \setminus A_t)$ are independent. Then, by the Claim,

$$\begin{aligned}\mathbb{P}(E_2) &= \mathbb{P}(N(A_t) = n) \cdot \mathbb{P}(N(A_{t+h} \setminus A_t) = 1) \\ &= p_n(t) \cdot (\mu(t+h) - \mu(t) + o(h)).\end{aligned}$$

Also,

$$\begin{aligned}q_n(t) &= \mathbb{P}(E) = \mathbb{P}(E_1) + \mathbb{P}(E_2) \\ &= q_n(t+h) + p_n(t) \cdot (\mu(t+h) - \mu(t) + o(h)).\end{aligned}$$

Hence,

$$q_n(t) - q_n(t+h) = p_n(t) \cdot (\mu(t+h) - \mu(t) + o(h)).$$

Dividing by h, and letting $h \to 0$, we obtain the differential equation:

$$-\frac{dq_n}{dt} = p_n \cdot \frac{d\mu}{dt}. \tag{1.4}$$

Given that $p_n = q_n - q_{n-1}$ and $p_0 = q_0 = \mathbb{P}(N(A_t) = 0)$, we can express p_1 as $p_1 = q_1 - q_0 = q_1 - p_0$, so $q_1 = p_1 + p_0$. For $n = 0$, we differentiate p_0:

$$-\frac{dp_0}{dt} = p_0 \cdot \frac{d\mu}{dt}. \tag{1.5}$$

Additionally, differentiating q_1:

$$-\frac{dq_1}{dt} = -\frac{d(p_1 + p_0)}{dt} = p_1 \cdot \frac{d\mu}{dt}.$$

From this, we obtain:

$$\frac{dp_1}{dt} = -\frac{dp_0}{dt} - p_1 \cdot \frac{d\mu}{dt} = (p_0 - p_1) \cdot \frac{d\mu}{dt}.$$

We now prove by induction that:

$$\frac{dp_n}{dt} = (p_{n-1} - p_n) \cdot \frac{d\mu}{dt}. \tag{1.6}$$

The case for $n = 1$ is already established in equation (1.5). Assuming the statement holds for $n - 1$, we proceed with the induction step for n:

$$\frac{dp_n}{dt} = \frac{d(q_n - q_{n-1})}{dt} = \frac{dq_n}{dt} - \frac{dq_{n-1}}{dt}.$$

Substituting by (1.4):

$$\frac{dp_n}{dt} = -p_n \cdot \frac{d\mu}{dt} + p_{n-1} \cdot \frac{d\mu}{dt} = (p_{n-1} - p_n) \cdot \frac{d\mu}{dt},$$

as desired.

Integrating equation (1.5) and using the initial condition $p_0(0) = 1$ and $\mu(0) = 0$, we obtain:

$$p_0(t) = e^{-\mu(t)}, \quad t \geq 0.$$

Similarly, from equation(1.6), we also have:

$$\begin{aligned}\frac{d(p_n \cdot e^{\mu})}{dt} &= \frac{dp_n}{dt} \cdot e^{\mu} + p_n \cdot \frac{d\mu}{dt} \\ &= (p_{n-1} - p_n) \cdot \frac{d\mu}{dt} \cdot e^{\mu} + p_n \cdot e^{\mu} \cdot \frac{d\mu}{dt} = p_{n-1} \cdot e^{\mu} \cdot \frac{d\mu}{dt}.\end{aligned}$$

Since $p_n(0) = 0$ for $n \geq 1$, we integrate to obtain:

$$p_n(t) = e^{-\mu(t)} \int_0^t p_{n-1}(s) \cdot e^{\mu(s)} \cdot \frac{d\mu}{ds} ds.$$

For $n = 1$, integrating yields:

$$\begin{aligned}p_1(t) &= e^{-\mu(t)} \int_0^t p_0(s) \cdot e^{\mu(s)} \frac{d\mu}{ds} ds = e^{-\mu(t)} \int_0^t \frac{d\mu}{ds} ds \\ &= e^{-\mu(t)} \cdot [\mu(t) - \mu(0)] = e^{-\mu(t)} \cdot \frac{\mu(t)^1}{1!}.\end{aligned}$$

Assuming the Lemma 1.7 holds for $n - 1$, substituting the variable $x = \mu(s)$, then, for n we get:

$$\begin{aligned}p_n(t) &= e^{-\mu(t)} \int_0^t \frac{(\mu(s))^{n-1}}{(n-1)!} \cdot e^{\mu(s)} \cdot \frac{d\mu}{ds} ds = \frac{e^{-\mu(t)}}{(n-1)!} \int_0^{\mu(t)} x^{n-1} dx \\ &= \frac{e^{-\mu(t)}}{(n-1)!} \cdot \frac{x^n}{n}\Big|_0^{\mu(t)} = e^{-\mu(t)} \cdot \frac{\mu(t)^n}{n!}.\end{aligned}$$

□

1.4 Poisson process in $\mathbb{R}$

When defined over the real line $\mathbb{R}$, the Poisson process models events occurring independently over continuous time. The homogeneous case has the property that the number of events in any bounded interval only depends on the length of the interval, not its location. This makes the Poisson process the natural starting point for studying continuous-time stochastic processes. In applications, Poisson processes on the line are especially crucial in queuing theory. They are used to model random arrivals of customers, calls, packets, or requests at a service facility.

So, consider the general case where the state space S is the real line $\mathbb{R}$. Assume that the measure $\mu(A)$ is finite for any bounded set A, (in the σ-algebra). In this scenario, the measure μ is completely determined by its values on intervals of the form $(a, b]$. To describe μ, we define a function $M(t)$ as follows:

$$M(t) = \begin{cases} \mu(0,t) = \mathbb{E}(N(0,t]) & \text{for } t \geq 0 \\ -\mu(t,0) = -\mathbb{E}(N(t,0]) & \text{for } t < 0. \end{cases}$$

The function M is an increasing function, and for any interval $(a, b]$ where $a < b$, the measure μ can be expressed as:

$$\mu(a,b] = M(b) - M(a).$$

This shows that μ is entirely determined by the function M, and is then called the *Stieltjes measure* associated with the increasing function M.
The measure μ is non-atomic if and only if M is continuous. The function M might be expressed as an integral of a rate function $\lambda(x)$:

$$M(t) = \int_0^t \lambda(x)dx.$$

If that is the case, then μ is given by $\mu(A) = \int_A \lambda(x)dx$.
A special case is the uniform Poisson process with a constant rate λ, where

$$M(t) = \lambda t, \quad t \in \mathbb{R}.$$

In this case, the measure μ is directly proportional to the length of the interval, reflecting a constant rate of occurrence across the real line. This uniform process is characterized by its invariance under translations, making it a fundamental example in the study of Poisson processes.

Chapter 2

Fundamental theorems of Poisson Processes

In this chapter we present four fundamental theorems in the study of Poisson point processes, which form the theoretical backbone for modeling random events in time and space. The superposition and mapping theorems are particularly valuable for constructing and transforming models, while the existence theorem justifies their use in continuous domains. The Bernoulli process remains essential for discrete-event scenarios.

2.1 Superposition Theorem

The Poisson process is characterized by several distinctive features that simplify its application and the determination of associated probabilities in a remarkably efficient manner. Among these features is the superposition property: if we combine the random points from independent Poisson processes, then we have a new Poisson process and the resulting intensity is just the sum of intensities of the combined processes.

However, it is important to emphasize that for this superposition to be valid, two conditions must be met: the original processes must be independent, and the total number of points in a region $N(A) = N_1(A)+N_2(A)$ must itself follow a Poisson distribution. This second condition is not automatic, it holds precisely because both $N_1(A)$ and $N_2(A)$ are Poisson and independent. Without these assumptions, we cannot conclude that the superposition results in a Poisson process.

The superposition property follows almost immediately from σ-additivity, except for a technical awkwardness, which is dealt with in the following

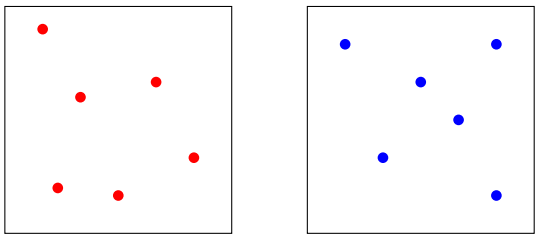


Figure 2.1: Π_1 and Π_2 independent Poisson processes.

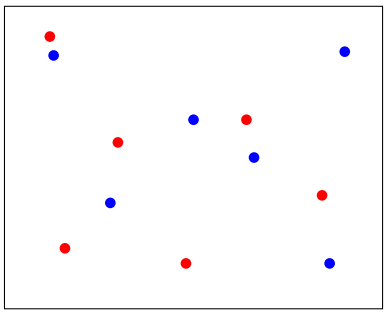


Figure 2.2: $\Pi_1 \cup \Pi_2$, superposition of Π_1 and Π_2.

lemma.

We start with a result that is a natural consequence of working with non-atomic measures, since the probability that two points appear at exactly the same location is zero.

Lemma 2.1 (**Disjointness**). *Assume that* Π_1 *and* Π_2 *are independent Poisson processes on the same state space* S *with mean measures* $\lambda_1(A)$ *and* $\lambda_2(A)$ *respectively. Let* A *be a measurable set with* $\mu_1(A)$, $\mu_2(A) < \infty$. *Then* Π_1 *and* Π_2 *are disjoint with probability 1 on* A*:*

$$\mathbb{P}(\Pi_1 \cap \Pi_2 \cap A = \emptyset) = 1. \tag{2.1}$$

The full proof of this lemma is technically involved (see Kingman's book Chapter 2 for details [4]), and we will omit it here, proceeding instead to discuss the superposition theorem.

Theorem 2.2 (**Superposition**). *Let* $\Pi_1, \Pi_2, \ldots$ *be a countable collection of independent Poisson processes defined on the same measurable space* S*. Let* Π_n *have mean measure* μ_n *for each* n*. Then the superposition*

$$\Pi = \cup_{n=1}^{\infty} \Pi_n$$

is a Poisson process with mean measure

$$\mu = \sum_{n=1}^{\infty} \mu_n. \tag{2.2}$$

Proof. Let $N_k(A)$ denote the number of points of Π_k in a measurable set A. Since $\Pi_1, \Pi_2, \ldots$ are independent Poisson processes, the random variables $N_k(A)$ are mutually independent, and

$$N(A) = \sum_{k=1}^{\infty} N_k(A) \tag{2.3}$$

is always true.

By the σ-additivity, the sum of independent Poisson random variables is also a Poisson random variable. Therefore, $N(A)$ follows a Poisson distribution with parameter $\mu(A)$ given by (2.2). This holds for any measurable set A with $\mu_k(A) < \infty$ for all k. If $\mu_k(A) = \infty$ for some k, then $N(A) = \infty$ almost surely, and the result (2.3) holds trivially.

To complete the proof, we need to show that for any collection of disjoint measurable sets $A_1, A_2, \ldots, A_m$, the random variables $N(A_1), N(A_2), \ldots,$ $N(A_m)$ are independent. This is clear because the variables in the double array

$$N_n(A_j), \quad n = 1, 2, \ldots, \quad j = 1, 2, \ldots, k$$

are all independent, and $N(A_j)$ is defined in terms of a subset of these variables disjoint from those for other j. □

2.2 Mapping Theorem

The key idea for the second great property of Poisson processes is that if a transformation is applied to the state space, the transformed points still follow a Poisson process, with the naturally transformed first intensity. [5]

However, a potential issue arises if the transformation causes multiple points to merge into the same location. This can affect the structure of the process.

Definition 2.3. *Let (S, Σ_1) and (T, Σ_2) be measurable spaces, and let $f : S \to T$ be a measurable mapping. For any measure μ on (S, Σ_1), the image of μ under f (also called the push-forward of μ) is the measure $f_*(\mu)$ defined by:*

$$f_*(\mu)(C) := \mu(f^{-1}(C)), \quad C \in \Sigma_2.$$

This means that the measure of a set C in Σ_2 under $f_*(\mu)$ is equal to the measure of its pre-image $f^{-1}(C)$ in Σ_1 under μ.

Let Π be a Poisson process on the state space S with mean measure μ. The points $f(x)$ for $x \in \Pi$ form a random countable set $f(\Pi) \subseteq T$. To determine whether $f(\Pi)$ is also a Poisson process, consider the number of points of $f(\Pi)$ in a measurable set $B \subseteq S$:

$$N^*(B) = \#\{f(\Pi) \cap B\}.$$

Assuming that the points $f(x)$ $(x \in \Pi)$ are distinct, this simplifies to:

$$N^*(B) = \#\{x \in \Pi;\ f(x) \in B\} = N(f^{-1}(B)),$$

where N is the counting measure associated with Π. Since $N(f^{-1}(B))$ has a Poisson distribution with mean $f_*(\mu)(B) = \mu(f^{-1}(B))$, it follows that $N^*(B)$ is distributed as $\mathcal{P}[\mu(f^{-1}(B))]$.

Moreover, if $B_1, B_2, \ldots, B_k$ are disjoint sets in Σ_2, their pre-images $f^{-1}(B_1)$, $f^{-1}(B_2), \ldots, f^{-1}(B_k)$ are also disjoint in S. Because Π is a Poisson process, the counts $N^*(B_j) = N(f^{-1}(B_j))$ are independent. Thus, $f(\Pi)$ satisfies the properties of a Poisson process on T with mean measure $f_*(\mu)$.

Thus $f(\Pi)$ is a Poisson process in T so long as the points $f(x)$ are distinct. This means that if two different points x_1 and x_2 from the Poisson process Π are mapped via f, their images $f(x_1)$ and $f(x_2)$ must not coincide (i.e., $f(x_1) \neq f(x_2)$). The new process's intensity measure, let's call it

$$\mu^*(B) := f_*(\mu)(B) = \mu(f^{-1}(B))$$

for $B \in \Sigma_2$, is derived from the original measure μ via f.

However, ensuring distinctness is not always trivial. For example, if f is a constant function, it maps all points in S to a single point in T, which is not a valid Poisson process.
More broadly, if the induced measure $f_*(\mu)$ has an atom (a point $t \in T$ with $f_*(\mu)(t) > 0$), then the pre-image set $A = f^{-1}(t)$ has positive measure under μ. This means there's a non-zero probability that two or more points from Π land in A and get mapped to the same t, breaking the distinctness requirement.

The key assumptions to avoid these problems are:

a) μ^* must be non-atomic.

b) μ should be σ-finite. Notably, μ^* doesn't need to be σ-finite.

We are finally ready to state the main result of this section.

Theorem 2.4 (**Mapping**). *Let Π be a Poisson process with σ-finite mean measure μ on the state space S, and let $f : S \to T$ be a measurable function such that the push-forward measure μ^* has no atoms. Then $f(\Pi)$ is a Poisson process on T having the induced measure μ^* as its mean measure.*

We restrict the proof to the case where f is injective, as the general case involves significant technical complexity. For a comprehensive treatment of the general case, we refer the reader to Kingman [4].

Proof. We assume that f is a measurable and injective function. Since f is measurable, for every measurable set $C \in \Sigma_2$, the pre-image

$$f^{-1}(C) = \{x \in S \mid f(x) \in C\}$$

is measurable in Σ_1.

The goal is to prove that the image process $f(\Pi) = \{f(x) \mid x \in \Pi\}$ is a Poisson process in T with intensity measure μ_*.

For a measurable set $C \in \Sigma_2$, define the function:

$$N^*(C) := \#\{f(\Pi) \cap C\},$$

which counts the number of points of $f(\Pi)$ that fall inside C.

Since f is injective, the points $f(x)$ for $x \in \Pi$ are distinct, so we can write:

$$N^*(C) = \#\{x \in \Pi \mid f(x) \in C\} = N(f^{-1}(C)),$$

where N is the counting measure associated with the original process Π.

Because Π is a Poisson process with intensity measure μ, we know that $N(f^{-1}(C)) \sim \mathcal{P}[\mu(f^{-1}(C))]$. By definition of the push-forward measure, this equals:

$$\mu(f^{-1}(C)) = f_*(\mu)(C) =: \mu^*(C).$$

Hence, $N^*(C) \sim \mathcal{P}(\mu^*(C))$.

Furthermore, if $C_1, C_2, \ldots, C_k \in \Sigma_2$ are disjoint measurable sets, then their pre-images $f^{-1}(C_1), \ldots, f^{-1}(C_k)$ are also disjoint in S. Since Π is a Poisson process, the counts $N(f^{-1}(C_j))$ are independent, and so the variables $N^*(C_j)$ are also independent.

We conclude that $f(\Pi)$ is a Poisson process on T with intensity measure $\mu^* = f_*(\mu)$. □

Example 2.5. To illustrate this theorem, consider a Poisson process Π in $\mathbb{R}^D$ with an intensity function $\lambda(x_1, x_2, \ldots, x_D)$. Our goal is to analyze what happens when we project Π onto a lower-dimensional space $\mathbb{R}^d$, where $d \leq D$. Consider the projection $f : \mathbb{R}^D \to \mathbb{R}^d$, defined as:

$$f(x_1, x_2, \ldots, x_D) = (x_1, x_2, \ldots, x_d).$$

Given a measurable set $B \subseteq \mathbb{R}^d$, the transformed mean measure $\mu^*(B)$ is obtained by integrating over the discarded dimensions:

$$\mu^*(B) = \int \cdots \int_{B \times \mathbb{R}^{D-d}} \lambda(x_1, x_2, \ldots, x_D)\, dx_1 dx_2 \ldots dx_D.$$

Rewriting this, we define a new intensity function $\lambda^*(x_1, x_2, \ldots, x_d)$, which captures the density of points in the reduced space $\mathbb{R}^d$:

$$\lambda^*(x_1, x_2, \ldots, x_d) = \int \cdots \int_{\mathbb{R}^{D-d}} \lambda(x_1, x_2, \ldots, x_D)\, dx_{d+1} dx_{d+2} \ldots dx_D. \quad (2.4)$$

Thus, the projected mean measure μ^* can now be rewritten as:

$$\mu^*(B) = \int \cdots \int_{B} \lambda^*(x_1, x_2, \ldots, x_d)\, dx_1 dx_2 \ldots dx_d.$$

Under the conditions that μ is σ-finite and μ^* is non-atomic, the projected process $f(\Pi)$ is also a Poisson process, now in $\mathbb{R}^d$, with the rate function λ^* if the integral (2.4) converges.

The result can be generalized to sets A that are countable unions of subsets where both μ_1 and μ_2 are finite. However, some form of finiteness is necessary, as illustrated by the following example.

Example 2.6. Consider a uniform Poisson process Π with rate $\lambda = 1$ on $\mathbb{R}^2$. Consider the projection $\phi(x, y) = x$ onto the first coordinate, so that the set

$$\phi(\Pi) = \{x | (x, y) \in \Pi\}$$

is a random countable subset of $\mathbb{R}$. If $A \subseteq \mathbb{R}$ is a Borel set, the number of points of $\phi(\Pi)$ that belong to it can be expressed as:

$$\#(\phi(\Pi) \cap A) = \#(\Pi \cap (A \times \mathbb{R})).$$

This quantify follows a Poisson distribution with mean given by the integral:

$$|A \times \mathbb{R}| = \int \int_{A \times \mathbb{R}} dx\ dy = \begin{cases} 0, & \text{if } |A| = 0 \\ +\infty, & \text{if } |A| > 0. \end{cases}$$

This implies that the random variable $\#(\phi(\Pi) \cap A)$ can only take the values 0 and ∞ with probability 1, meaning that it follows the distribution $\mathcal{P}(\mu)$ with $\mu = 0$ or $\mu = \infty$. Therefore, $\phi(\Pi)$ is itself a Poisson process on $\mathbb{R}$, with degenerate mean measure:

$$\mu(A) = \begin{cases} 0, & \text{if } |A| = 0, \\ \infty, & \text{if } |A| > 0. \end{cases}$$

Now, consider two independent Poisson processes, Π_1 and Π_2, defined as copies of $\phi(\Pi)$. Since the counts $N_1(A_1)$ and $N_2(A_2)$ of these processes are completely degenerate (i.e., they take trivial values), they are automatically independent. This means that while Π_1 and Π_2 behave like independent Poisson processes, the expected relation in equation (2.1) no longer holds when $|A| > 0$.

One possible objection is that Π_1 and Π_2 are not truly independent, but only appear to be because they are described in terms of their own point sets. If we were to use a more refined measurable structure for infinite point sets, we would see that Π_1 and Π_2 are simply duplicates of the same random set $\phi(\Pi)$.

To conclude this section, we present a result that is almost too straightforward to state explicitly. Its proof simply involves verifying the definition, but it is used so frequently that it deserves to be highlighted.

Theorem 2.7 (**Restriction**). *Let Π be a Poisson process on a measurable space (S, Σ, μ) with intensity measure μ. If $S_1 \subseteq S$ is a measurable subset of S, the the restriction of Π to S_1, denoted Π_1, is also a Poisson process with intensity measure $\mu|_{S_1}$, the restriction of μ to A.*

In other words, the random countable set Π can be interpreted in two ways:

1. As a Poisson process on S with mean measure
$$\mu_1(A) = \mu(A \cap S_1).$$
2. As a Poisson process on S_1 whose mean measure is the restriction of μ to S_1.

2.3 Conditioning. The Bernoulli process

Consider a Poisson process Π defined on a space S with a finite mean measure μ (i.e., $\mu(S) < \infty$). In this case, Π is almost surely a finite subset of S, and the total number of points $N(S)$ follows a Poisson distribution $\mathcal{P}(\mu(S))$. Now, imagine we condition Π on the event that $N(S) = n$. What does the process look like under this condition? We will show that conditioning a Poisson process on a fixed total number of points turns it into a binomial/multinomial point process.

We have introduced the *binomial distribution* in Chapter 1, Section 1.2. Let us recall what is a *multinomial distribution.* [6]

The multinomial distribution arises from an experiment with the following properties:

- a fixed number n of trials, where each trial is independent of the others,
- each trial has k mutually exclusive and exhaustive possible outcomes, denoted by $A_1, A_2, \ldots, A_k$,
- on each trial, A_j occurs with probability $p_j, \quad j = 1, \ldots, k$.

If we let N_j count the number of trials for which outcome A_j occurs, then the random vector $N = (N_1, N_2, \ldots, N_k)$ is said to have a *multinomial distribution with index n and parameter vector $p = (p_1, \ldots, p_k)$*, which we denote as

$$N \sim \mathit{Mult}(n, p).$$

Note that we must have $p_1 + p_2 + \cdots + p_k = 1$ and $N_1 + N_2 + \cdots + N_k = n$. Its distribution is given, for $0 \leq n_1, \ldots, n_k \leq N$, by

$$\mathbb{P}(N_1 = n_1, \ldots, N_k = n_k) = \frac{n!}{n_1! n_2! \ldots n_k!} p_i^{n_1} p_2^{n_2} \ldots p_k^{n_k}. \tag{2.5}$$

This generalizes the binomial distribution ($k = 2$) to multiple categories.

Returning to what we want to show, let $\mathbb{P}_n$ represent the conditional probability measure:

$$\mathbb{P}_n\{\cdot\} = \mathbb{P}\{\cdot | N(S) = n\}.$$

Let $A_1, A_2, \ldots, A_k$ be disjoint subsets of S and let $n_1, n_2, \ldots, n_k$ be non-negative integers. The aim is to determine the probability

$$\begin{aligned} &\mathbb{P}\{N(A_1) = n_1, \ldots, N(A_k) = n_k | N(S) = n\} \\ &\qquad = \mathbb{P}_n\{N(A_1) = n_1, \ldots, N(A_k) = n_k\}. \end{aligned}$$

Consider $A_0 = (A_1 \cup A_2 \cup \cdots \cup A_k)^c$ and $n_0 = n - \sum_{j=1}^k n_j$. By the independence of the different $N(A_j)$,

$$\begin{aligned}
\mathbb{P}_n\{N(A_1) = n_1, \ldots, N(A_k) = n_k\} &= \mathbb{P}\{N(A_1) = n_1, \ldots, N(A_k) = n_k \mid N(S) = n\} \\
&= \frac{\mathbb{P}\{[N(A_1) = n_1, \ldots, N(A_k) = n_k] \cap [N(S) = n]\}}{\mathbb{P}\{N(S) = n\}} \\
&= \frac{\mathbb{P}\{N(A_1) = n_1, \ldots, N(A_k) = n_k, N(A_0) = n_0\}}{\mathbb{P}\{N(S) = n\}} \\
&= \frac{\prod_{j=0}^k \mathbb{P}(N(A_j) = n_j)}{\mathbb{P}(N(S) = n)}.
\end{aligned}$$

Since $N(A_j)$ is a Poisson variable of parameter $\mu(A_j)$,

$$\begin{aligned}
\mathbb{P}_n\{N(A_1) = n_1, \ldots, N(A_k) = n_k\} &= \frac{\prod_{j=0}^k \frac{e^{-\mu(A_j)}\mu(A_j)^{n_j}}{n_j!}}{\frac{e^{-\mu(S)}\mu(S)^n}{n!}} \\
&= \frac{n!}{n_0! n_1! \cdots n_k!} \left(\frac{\mu(A_0)}{\mu(S)}\right)^{n_0} \left(\frac{\mu(A_1)}{\mu(S)}\right)^{n_1} \cdots \left(\frac{\mu(A_k)}{\mu(S)}\right)^{n_k}.
\end{aligned}$$

Letting

$$p_j = p(A_j) = \frac{\mu(A_j)}{\mu(S)} \tag{2.6}$$

we see thus that for $\Pi \subseteq S$ with $\#\Pi = n$

$$\mathbb{P}\{N(A_1) = n_1, \ldots, N(A_k) = n_k\} = \frac{n!}{n_0! n_1! \cdots n_k!} p(A_0)^{n_0} p(A_1)^{n_1} \cdots p(A_k)^{n_k},$$

that is $N = (N(A_1), \ldots, N(A_k))$ is a Bernoulli process of parameter vector $p = (p_1, \ldots, p_k)$.

There's a simpler method to build a Bernoulli process using independent random variables $X_1, X_2, \ldots, X_n$, each distributed according to a probability measure μ. If μ has no atoms (no single point has positive probability), then the X_i values are almost surely distinct. This means the set $\{X_1, X_2, \ldots, X_n\}$ is random with exactly n unique elements with probability one. It is then an easy multinomial calculation that the counts

$$N_r(A) = \#\{r : X_r \in A\}$$

satisfy (2.5).

Hence, for a finite Poisson process, given the total count $N(S)$, the points behave like $N(S)$ independent random variables with a common distribution (2.6). This observation has important implications, one of which is explored further in the next section.

2.4 The Existence Theorem

Given a non-atomic and σ-finite measure μ on a space S, a natural question arises: Does there exist a Poisson process with intensity μ? Such a process would satisfy that $N(A)$ has distribution $\mathcal{P}(\mu(A))$ for all measurable subset A of S, and that $N(A)$ is independent of $N(B)$ if $A \cap B = \emptyset$.

The answer is yes, and the construction is possible under very general conditions on μ. This result ensures that Poisson processes can be rigorously defined on a wide variety of spaces, leading to a highly flexible and broadly applicable theory.

Theorem 2.8 (**Existence**). *Let μ be a non-atomic and σ-finite measure on S that can be decomposed as $\mu = \sum_n \mu_n$ where $\mu_n(S) < \infty$ for all n. Define N_n and $X_{n,j}$, $n, j \geq 1$ independent random variables with*

$$N_n \sim \mathcal{P}(\mu_n(A)), \quad X_{n,j} \sim p_n,$$

where $p_n(\cdot) = \frac{\mu_n(\cdot)}{\mu_n(S)}$. Then the process

$$\Pi := \bigcup_n \{X_{n,1}, \ldots, X_{n,N_n}\}$$

is a Poisson process with mean measure μ.

Proof. First we prove that $\Pi_n := \{X_{n,1}, \ldots, X_{n,N_n}\}$ is a Poisson process. As shown in Section 2.3, conditioning to $N_n = m$, the process is Bernoulli:

$$\begin{aligned} P(N_n(A_1) = m_1, \ldots, N_n(A_k) = m_k \mid N_n = m) \\ = \frac{m!}{m_0! \ldots m_k!} \cdot \left(\frac{\mu_n(A_0)}{\mu_n(S)}\right)^{m_0} \cdots \left(\frac{\mu_n(A_k)}{\mu_n(S)}\right)^{m_k} \end{aligned}$$

where $m_0 = m - (m_1 + \cdots + m_k)$ and $A_0 = S \setminus (A_1 \cup \cdots \cup A_k)$. Multiplying both terms by $P(N_n(S) = m) = e^{-\mu_n(S)}(\mu_n(S))^m/m!$ and summing over m,

we get

$$\begin{aligned}
&P(N_n(A_1) = m_1, \ldots, N_n(A_k) = m_k) \\
&\quad = \sum_{m \geq 0} \frac{e^{-\mu_n(S)}(\mu_n(S))^m}{m!} \cdot \frac{m!}{m_0! \cdots m_k!} \cdot \left(\frac{\mu_n(A_0)}{\mu_n(S)}\right)^{m_0} \cdots \left(\frac{\mu_n(A_k)}{\mu_n(S)}\right)^{m_k} \\
&\quad = \sum_{m \geq 0} \frac{e^{-\mu_n(A_0)}(\mu_n(A_0))^{m_0}}{m_0!} \cdots \frac{e^{-\mu_n(A_k)}(\mu_n(A_k))^{m_k}}{m_k!} \\
&\quad = \prod_{i=1}^{k} \frac{e^{-\mu_n(A_i)}(\mu_n(A_i))^{m_i}}{m_i!}.
\end{aligned}$$

Thus, the $N_n(A_j)$ are independent random variables with distributions $\mathcal{P}(\mu_n(A_j))$. Hence, Π_n is a Poisson process.

Since Π_n are independent Poisson processes with respective intensities μ_n, the superposition Theorem implies that Π defines a Poisson process with mean measure μ. □

Chapter 3

Linear statistics. Campbell's theorem

This chapter introduces three foundational results in the analysis of Poisson point processes, forming the basis for modeling random structures in continuous spaces. Linear statistics offer a systematic way to analyze sums over random configurations, Campbell's Theorem provides exact formulas for their moments, and Rényi's characterization links Poisson counts to spatial independence, showing how local behavior determines global structure.

3.1 Linear statistics

Given a point process Π and a measurable function f, the quantity

$$\Lambda_\Pi(f) = \sum_{x \in \Pi} f(x) \tag{3.1}$$

is sometimes called the "linear statistic" associated with f. Note that

$$\Lambda_\Pi(\alpha f + \beta g) = \alpha \Lambda_\Pi(f) + \beta \Lambda_\Pi(g) \quad \text{for } \alpha, \beta \in \mathbb{R},$$

which shows that Λ_Π is a linear function in f. This quantity, when varied over f, provides diverse information about Π. For example, if $f = \mathcal{X}_A$, then $\Lambda_\Pi(A) = N(A)$.

The goal is to understand two main aspects: the conditions under which this sum converges and the main properties of Λ_Π, such as its expectation, variance and distribution. We will focus on presenting the essential ideas and intuitions, rather than formal mathematical proofs.

Sums of the form (3.1) arise in a wide range of applications involving Poisson processes.

In temporal settings, a common example is the *shot effect*, where each point in a one-dimensional Poisson process produces an effect that persists over time. For instance, particle emissions may generate a decaying influence modeled by a function like $\phi(t - T_j)$, where T_j denotes the time of the event. In spatial contexts, a typical application involves computing the gravitational potential generated by stars randomly distributed in space according to a Poisson process.

The method for determining the distributions of Λ_Π involves first computing it for simple functions f and then generalizing the result to more complex functions using standard techniques from integration theory.

Begin by considering a simple function $f = \sum_{j=1}^{k} f_j \mathcal{X}_{A_j}$, with

$$A_j = \{x \in S \mid f(x) = f_j\}$$

measurable sets and $\mu(A_j) < \infty$. The sum Λ_Π can then be expressed as:

$$\Lambda_\Pi = \sum_{j=1}^{k} f_j N(A_j).$$

This representation fully determines the distribution of Λ_Π, most conveniently characterized by its moment generating function. By (1.2), for any real or complex θ, the moment generating function of Λ_Π is given by:

$$\mathbb{E}(e^{\theta\Lambda_\Pi}) = \prod_{j=1}^{k} \mathbb{E}(e^{\theta f_j N(A_j)}) = \prod_{j=1}^{k} \exp\left\{(e^{\theta f_j} - 1)\mu(A_j)\right\}.$$

Using the fact that $f(x) = f_j$ on A_j, this can be written as:

$$\mathbb{E}(e^{\theta\Lambda_\Pi}) = \exp\left\{\sum_{j=1}^{k} \int_{A_j} (e^{\theta f(x)} - 1) d\mu(x)\right\}.$$

Since $f(x) = 0$ outside the union of the sets A_j, the integral extends over the entire space S:

$$\mathbb{E}(e^{\theta\Lambda_\Pi}) = \exp\left\{\int_S (e^{\theta f(x)} - 1) d\mu(x)\right\}. \tag{3.2}$$

This is the master equation governing the distribution of Λ_Π.

There are some observations about its convergence and some special cases. If f takes both positive and negative values, it is natural to let $\theta = it$. This yields the characteristic function of Λ_Π:

$$\mathbb{E}(e^{it\Lambda_\Pi}) = \exp\left\{\int_S (e^{itf(x)} - 1) d\mu(x)\right\}. \tag{3.3}$$

Knowing this for all real t fully determines the distribution of Λ_Π.
The sum $\Lambda_\Pi = \sum_{j=1}^{k} f_j N(A_j)$ converges absolutely with probability 1 if and only if the integral in (3.3) converges absolutely for some real $t \neq 0$.
If f takes only positive values, it is often more convenient to work with negative real $\theta = -u, u > 0$, so that (3.2) becomes

$$\mathbb{E}(e^{-u\Lambda_\Pi}) = \exp\left\{-\int_S \left(1 - e^{-uf(x)}\right) d\mu(x)\right\}.$$

Expanding the master equation (3.2) as a power series in θ and equating coefficients yields the first two moments of Λ_Π.
We begin by analyzing the left-hand side, focusing on the moment-generating function $\mathbb{E}(e^{\theta\Lambda_\Pi})$. To expand this, we use a Taylor series expansion of the exponential function $e^{\theta\Lambda_\Pi}$ around $\theta = 0$:

$$e^{\theta\Lambda_\Pi} = 1 + \theta\Lambda_\Pi + \frac{(\theta\Lambda_\Pi)^2}{2!} + \mathcal{O}(\theta^3).$$

Next, we apply the expectation operator to both sides. By the linearity, we obtain:

$$\mathbb{E}(e^{\theta\Lambda_\Pi}) = 1 + \theta\mathbb{E}(\Lambda_\Pi) + \frac{\theta^2}{2}\mathbb{E}(\Lambda_\Pi^2) + \mathbb{E}(\mathcal{O}(\theta^3)).$$

Moving to the right-hand side, we begin with the integral term. Using a Taylor expansion of $e^{\theta f(x)} - 1$ around $\theta = 0$ and integrating term by term over the space S with respect to the measure μ, we obtain:

$$\int_S (e^{\theta f(x)} - 1)\, d\mu(x) = \theta \int_S f(x)\, d\mu(x) + \frac{\theta^2}{2} \int_S f(x)^2\, d\mu(x) + \mathcal{O}(\theta^3).$$

For conciseness, let:

$$A = \int_S f(x) d\mu(x), \quad B = \int_S f(x)^2 d\mu(x).$$

This simplifies the integral to:

$$\int_S (e^{\theta f(x)} - 1) d\mu(x) = \theta A + \frac{\theta^2}{2} B + \mathcal{O}(\theta^3).$$

Let $y = \theta A + \frac{\theta^2}{2}B + \mathcal{O}(\theta^3)$. Using the Taylor expansion of e^y around $y = 0$ and keeping terms up to θ^2, we have:

$$e^y = 1 + (\theta A + \frac{\theta^2}{2}B) + \frac{(\theta A)^2}{2} + \mathcal{O}(\theta^3) = 1 + \theta A + \frac{\theta^2}{2}(B + A^2) + \mathcal{O}(\theta^3).$$

Finally, substituting back A and B, we arrive at the expansion of the exponential integral:

$$\begin{aligned}\exp\left\{\int_S \left(e^{\theta f(x)} - 1\right) d\mu(x)\right\} = 1 &+ \theta \int_S f(x)\, d\mu(x) \\ &+ \frac{\theta^2}{2}\left(\int_S f(x)^2\, d\mu(x) + \left(\int_S f(x)\, d\mu(x)\right)^2\right) \\ &+ \mathcal{O}(\theta^3).\end{aligned}$$

By comparing the first-order terms in the two expansions, we obtain the expectation of Λ_Π:

$$\mathbb{E}(\Lambda_\Pi) = \int_S f(x) d\mu(x). \tag{3.4}$$

For the second moment, we compare the second-order terms and find:

$$\mathbb{E}(\Lambda_\Pi^2) = \int_S f(x)^2 d\mu(x) + (\int_S f(x) d\mu(x))^2.$$

Thus, the variance of Λ_Π is derived by:

$$\mathbb{V}(\Lambda_\Pi) = \mathbb{E}(\Lambda_\Pi^2) - \mathbb{E}(\Lambda_\Pi)^2 = \int_S f(x)^2 d\mu(x). \tag{3.5}$$

For two real-valued functions f_1 and f_2, replace θf with $\theta_1 f_1 + \theta_2 f_2$ in (3.2) to obtain the joint moment generating function:

$$\mathbb{E}\left(e^{\theta_1\Lambda_1 + \theta_2\Lambda_2}\right) = \exp\left\{\int_S \left(e^{\theta_1 f_1(x) + \theta_2 f_2(x)} - 1\right) d\mu(x)\right\} \tag{3.6}$$

where:

$$\Lambda_1 = \sum_{X\in\Pi} f_1(X), \quad \Lambda_2 = \sum_{X\in\Pi} f_2(X).$$

This result generalizes naturally to three or more functions.
By a similar computation to the one above, the coefficient of $\theta_1\theta_2$ in the power series of expansion (3.6) yields the cross-moment:

$$\mathbb{E}(\Lambda_1\Lambda_2) = \left(\int_S f_1 d\mu\right)\left(\int_S f_2 d\mu\right) + \int_S f_1 \cdot f_2\ d\mu. \tag{3.7}$$

This implies, in particular, that

$$cov(\Lambda_1, \Lambda_2) = \mathbb{E}(\Lambda_1\Lambda_2) - \mathbb{E}(\Lambda_1)\mathbb{E}(\Lambda_2) = \int_S f_1(x)f_2(x)d\mu(x).$$

The left-hand side of (3.7) can be decomposed as:

$$\mathbb{E}(\sum_{X_1,X_2\in\Pi} f_1(X_1)f_2(X_2)) = \mathbb{E}(\sum_{X\in\Pi} f_1(X)f_2(X)) + \mathbb{E}(\sum_{X_1\neq X_2} f_1(X_1)f_2(X_2)).$$

The first term on the right is $\int f_1 f_2 d\mu$, so (3.7) leads to the important identity:

$$\begin{aligned}\mathbb{E}\left(\sum_{\substack{X_1,X_2\in\Pi\\ X_1\neq X_2}} f_1(X_1)f_2(X_2)\right) &= \left(\int_S f_1(x)d\mu(x)\right)\left(\int_S f_2(x)d\mu(x)\right)\\ &= \mathbb{E}\left(\sum_{X_1\in\Pi} f_1(X_1)\right)\mathbb{E}\left(\sum_{X_2\in\Pi} f_2(X_2)\right).\end{aligned}$$

This result extends naturally to products of n functions:

$$\mathbb{E}\left(\sum_{X_1,\ldots,X_n\in\Pi \text{ distinct}} \prod_{j=1}^{n} f_j(X_j)\right) = \prod_{j=1}^{n}\mathbb{E}\left(\sum_{X_j\in\Pi} f_j(X_j)\right) = \prod_{j=1}^{n}\mathbb{E}(\Lambda_j).$$

These fundamental results originate from N. R. Campbell's pioneering work and are collectively known as Campbell's Theorem. They provide powerful tools for analyzing distributions arising from Poisson processes.

3.2 Campbell's Theorem

The rigorous derivations in the previous section are validated by first applying identity (3.2) to simple functions f and then generalizing it to broader function classes via integration theory. These findings are consolidated in Campbell's Theorem. [8]

Theorem 3.1 (**Campbell's Theorem**). *For a Poisson point process* Π *on* S *with intensity measure* μ *and a measurable function* $f : S \to \mathbb{R}$,

a) The random sum

$$\Lambda = \sum_{X\in\Pi} f(X)$$

is absolutely convergent with probability one if and only if

$$\int_S \min(|f(x)|, 1)d\mu(x) < \infty. \tag{3.8}$$

b) Provided that (3.8) is satisfied, for any complex value θ the equation

$$\mathbb{E}(e^{\theta\Lambda}) = \exp(\int_S (e^{\theta\Lambda} - 1)d\mu(x) \tag{3.9}$$

holds as long as the integral on the right-hand side converges, which is the case for purely imaginary θ. Moreover,

$$\mathbb{E}(\Lambda) = \int_S f(x)d\mu(x),$$

in the sense that the expectation exists if and only if the integral converges, and they are then equal. Then, also

$$Var(\Lambda) = \int_S f(x)^2 d\mu(x).$$

Remark 3.2. It is an easy computation that for $y \in \mathbb{R}$

$$|e^{iy} - 1| \leq 2\min(|y|, 1).$$

This shows that if

$$\int_S \min(|f(x)|, 1)\, d\mu(x) < \infty,$$

then for any fixed $\theta \in i\mathbb{R}$ (pure imaginary),

$$\int_S |e^{\theta f(x)} - 1|\, d\mu(x) < \infty.$$

Proof. Equation (3.9) is already verified for all complex θ when f is simple, as seen in the previous section.

The argument follows the standard approach of measure-theoretic integration: we first establish the result for characteristic functions, then extend it to simple functions, and finally generalize to arbitrary measurable functions.

Step 1: Let $E \subseteq S$ be a measurable set with finite measure, $\mu(E) < +\infty$ and consider the characteristic function, given by:

$$\chi_E = \begin{cases} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E. \end{cases}$$

In this case,

$$\Lambda = \sum_{x \in \Pi} \chi_E(x) = N(E).$$

Let us check the validity of equation (3.9). On the right-hand side we have:

$$\exp\left(\int_S (e^{\theta\chi_E} - 1)d\mu(x)\right) = \exp\left(\int_E (e^{\theta} - 1)d\mu(x)\right) = \exp\left((e^{\theta} - 1)\mu(E)\right).$$

Since $\Lambda = N(E)$, the left-hand side can be written as follows:

$$\begin{aligned}\mathbb{E}(e^{\theta\Lambda}) = \mathbb{E}(e^{\theta N(E)}) &= \sum_{k=0}^{\infty} e^{\theta k} \cdot P(N(E) = k) \\ &= e^{-\mu(E)} \sum_{k=0}^{\infty} \frac{(e^{\theta}\mu(E))^k}{k!} = e^{-\mu(E)} \exp(e^{\theta}\mu(E)) = e^{\mu(E)(e^{\theta}-1)}.\end{aligned}$$

Step 2: Assume s is a simple function, $s = \sum_{j=1}^{N} \alpha_j \chi_{E_j}$ where $\{\alpha_j\}_{j=1}^{N}$ are real coefficients and the measurable sets $\{E_j\}_{j=1}^{N}$ are pairwise disjoint. Thus, f is given by:

$$s(x) = \begin{cases} \alpha_j & \text{if } x \in E_j \\ 0 & \text{if } x \notin \cup_{j=1}^{N} E_j. \end{cases}$$

Let's check now (3.9), as in the previous case. On the right-hand side:

$$\begin{aligned}\exp\left(\int_S (e^{\theta s(x)} - 1)d\mu(x)\right) &= \exp\left(\int_{\cup_j E_j} (e^{\theta s(x)} - 1)d\mu(x)\right) \\ &= \exp\left(\sum_{j=1}^{N} \int_{E_j} (e^{\theta\alpha_j} - 1)d\mu(x)\right) \\ &= \exp\left(\sum_{j=1}^{N} (e^{\theta\alpha_j} - 1)\mu(E_j)\right) = \prod_{j=1}^{N} e^{(e^{\theta\alpha_j}-1)\mu(E_j)}.\end{aligned}$$

Using the identity $\Lambda = \sum_{X\in\Pi} f(X) = \alpha_1 N(E_1) + \cdots + \alpha_N N(E_N)$ and recalling that the measurable sets $\{E_j\}$ are disjoint (hence the $N(E_j)$ are independent), we see that the other side of the equation yields the same result:

$$\begin{aligned}\mathbb{E}(e^{\theta\Lambda}) &= \mathbb{E}(e^{\theta(\alpha_1 N(E_1) + \cdots + \alpha_N N(E_N))}) \\ &= \prod_{j=1}^{N} \mathbb{E}(e^{\theta\alpha_j N(E_j)}) = \prod_{j=1}^{N} e^{(e^{\theta\alpha_j}-1)\mu(E_j)}.\end{aligned}$$

Step 3: Having established equation (3.9) for simple functions, we now extend the argument to $f \geq 0$. Consider a sequence of simple functions $\{S_m\}$ such that $0 \leq S_m \leq f$ and $f(x) = \lim_m S_m(x)$ for all $x \in S$.

We aim to take the limit of the equality:

$$\mathbb{E}(e^{\theta\Lambda_m}) = \exp\left(\int_S (e^{\theta S_m(x)} - 1)d\mu(x)\right),$$

where$\Lambda_m = \sum_{\lambda\in\Pi} S_m(\lambda)$.)
To avoid convergence issues with the integrals, we begin with $\theta \in \mathbb{R}^-$, setting $\theta = -u$, with $u > 0$.
On the left side we apply the dominated convergence theorem (see Annex A for the statement). Thus:

$$\lim_{m\to\infty} \mathbb{E}(e^{-u\Lambda_m}) = \lim_{m\to\infty} \int_\Omega e^{-u\Lambda_m} d\mathbb{P} = \int_\Omega e^{-u\Lambda} d\mathbb{P} = \mathbb{E}(e^{-u\Lambda}),$$

since $|e^{-u\Lambda_m}| \leq 1$ and 1 is an integrable function with respect to $\mathbb{P}$.
On the other side of (3.9), we apply the monotone convergence theorem (see Annex A). The functions $g_m(x) = 1 - e^{-uf_m(x)}$ are increasing in m; therefore,

$$\begin{aligned}\lim_{m\to\infty} \int_S (1 - e^{-uS_m(x)})d\mu(x) &= \int_S (1 - e^{-uf(x)})d\mu(x) \\ &= -\int_S (e^{-uf(x)} - 1)d\mu(x).\end{aligned}$$

Thus

$$\lim_{m\to\infty} \exp\left(\int_S (e^{-uS_m(x)} - 1)d\mu(x)\right) = \exp\left(\int_S (e^{-uf(x)} - 1)d\mu(x)\right)$$

and we deduce the following equality: for $f \geq 0$ and $u > 0$,

$$\mathbb{E}(e^{-u\Lambda}) = \exp\left(\int_S (e^{-uf(x)} - 1)\, d\mu(x)\right). \tag{3.10}$$

We now examine the limiting behavior as $u \to 0$. It is evident that, pointwise

$$\lim_{u\to 0} e^{-uf(x)} - 1 = 0.$$

We now bound $|e^{-uf(x)} - 1|$ with a function $g(x)$ such that

$$\int_S g(x)\, d\mu(x) < \infty$$

in order to apply the dominated convergence theorem and pass the pointwise limit inside the integral.
If $f(x) \geq 1$, then we bound:

$$|e^{-uf(x)} - 1| = 1 - e^{-uf} \leq 1.$$

If $f(x) < 1$, we use the bound $1 - e^{-y} \leq y$, when $y \in (0,1)$, with $y = uf(x)$, to get, for u near 0

$$|e^{-uf(x)} - 1| = 1 - e^{-uf(x)} \leq uf(x) \leq f(x) = |f(x)|.$$

Thus, all combined

$$|e^{-uf(x)} - 1| \leq \min(1, f(x)) = g(x),$$

which, by hypothesis, is integrable in S.
Therefore, by the dominated convergence theorem:

$$\begin{aligned}\lim_{u\to 0} \mathbb{E}(e^{-u\Lambda}) &= \lim_{u\to 0} \exp\left(\int_S \left(e^{-uf(x)} - 1\right)\, d\mu(x)\right) \\ &= \exp\left(\int_S (e^0 - 1)\, d\mu(x)\right) = 1.\end{aligned}$$

Let us observe that this implies that $\Lambda = \sum_{\lambda\in\Pi} f(\lambda)$ must be finite with probability 1. Suppose that this is not the case, that is, $\Lambda = +\infty$ for all $\omega \in A,\ A \subseteq \Omega$, with $\mathbb{P}(A) > 0$. Then we would have:

$$\begin{aligned}\mathbb{E}(e^{-u\Lambda}) &= \int_\Omega e^{-u\Lambda}\, d\mathbb{P} = \int_A e^{-\infty}\, d\mathbb{P} + \int_{A^c} e^{-u\Lambda}\, d\mathbb{P} \\ &= \int_{A^c} e^{-u\Lambda}\, d\mathbb{P} \leq \int_{A^c} 1\, d\mathbb{P} = \mathbb{P}(A^c) = 1 - \mathbb{P}(A) < 1,\end{aligned}$$

and therefore also

$$\lim_{u\to 0} \mathbb{E}(e^{-u\Lambda}) \leq 1 - \mathbb{P}(A) < 1,$$

a contradiction what we have just seen.
This proves that, if equation (3.8) holds, then the equality holds for all $u < 0$, or equivalently, it holds for all $\theta \in \mathbb{C}$ such that $\mathrm{Re}(\theta) \leq 0$. Observe that for such θ, we can write $\theta = -u + iv$, with $u \geq 0$, and $e^{\theta\Lambda} = e^{-u\Lambda}e^{iv\Lambda}$. That is, we multiply $e^{-u\Lambda}$ by a factor of modulus 1, which does not affect the absolute convergence.

Since both sides of equation (3.9) are well-defined and analytic in θ, and since they coincide on $\mathbb{R}^-$, they must be the same function. Thus, the theorem is proved if $f \geq 0$.

It is straightforward to prove that

$$e^{-1}\min(1, y) \leq 1 - e^{-y} \leq \min(1, y)$$

for all $y > 0$.

This proves, by comparison, that

$$\begin{aligned}\int_S (1-e^{-uf(x)})d\mu(x) < \infty &\iff \int_S \min(1, uf(x))d\mu(x) < \infty \\ &\iff u \cdot \int_S \min(1, f(x))d\mu(x) < \infty \\ &\iff \int_S \min(1, f(x))d\mu(x) < \infty.\end{aligned}$$

In particular, regarding equality (3.10), if condition (3.8) fails to hold, it follows that $\int_S (e^{-uf(x)} - 1)d\mu(x) = -\infty$ and $\mathbb{E}(e^{-u\Lambda}) = 0$, which implies $\Lambda = +\infty$ with probability 1.

Step 4: To prove the general case, where f can be any function, write:

$$f^+ = \max(f, 0), \quad f^- = \max(-f, 0),$$

so that $f = f^+ - f^-$.
The sum $\Lambda(f) = \Lambda(f^+) - \Lambda(f^-)$ converges absolutely if and only if

$$\Lambda_+ = \Lambda(f^+) = \sum_{X \in \Pi} f^+(X) = \sum_{X \in \Pi_+} f(X)$$

and

$$\Lambda_- = \Lambda(f^-) = \sum_{X \in \Pi} f^-(X) = -\sum_{X \in \Pi_-} f(X)$$

converge, where Π_+ and Π_- are the restrictions of Π to

$$S_+ = \{f > 0\}, \quad S_- = \{f < 0\}.$$

This shows that (3.8) is the necessary and sufficient condition for convergence. If it holds, and θ is pure imaginary, then, once more by the independence

$$\begin{aligned}\mathbb{E}(e^{\theta\Lambda}) &= \mathbb{E}(e^{\theta\Lambda_+ - \theta\Lambda_-}) \\ &= \mathbb{E}(e^{\theta\Lambda_+})\mathbb{E}(e^{-\theta\Lambda_-}) \\ &= \exp\left(\int_S (e^{\theta f^+} - 1)\, d\mu(x)\right) \exp\left(\int_S (e^{-\theta f^-} - 1)\, d\mu(x)\right) \\ &= \exp\left(\int_S \left(e^{\theta f^+} + e^{-\theta f^-} - 2\right)\, d\mu(x)\right) \\ &= \exp\left(\int_S (e^{\theta f} - 1)\, d\mu(x)\right).\end{aligned}$$

Thus (3.9) holds at least for imaginary θ and its extension to all complex θ for which the integral converges is a matter of analytic continuation. □

Corollary 3.3. *Let $f_1, f_2, \ldots, f_n$ be measurable functions satisfying identity (3.9) (guaranteeing that the random sums $\Lambda_j = \sum_{X\in\Pi} f_j(X), \quad j = 1, \ldots, n$ converge with probability 1). Then, the joint characteristic function of these sums is given by:*

$$\mathbb{E}\left[\exp\left(i\sum_{j=1}^{n} t_j\Lambda_j\right)\right] = \exp\left(\int_S \left(\exp\left(i\sum_{j=1}^{n} t_j f_j(x)\right) - 1\right) d\mu(x)\right). \quad (3.11)$$

Furthermore, if each f_j is square-integrable with respect to μ, that is,

$$\int_S f_j(x)^2\mu(dx) < \infty, \ j = 1, \ldots, n,$$

then the covariance between any two sums Λ_j and Λ_k is:

$$cov(\Lambda_j, \Lambda_k) = \int_S f_j(x)f_k(x)\mu(dx).$$

Proof. The joint characteristic function (3.11) is derived by substituting the linear combination $f = t_1f_1 + \cdots + t_nf_n$ into the earlier result (3.9), with $\theta = i$. Condition (3.8) ensures that $t_1f_1 + \cdots + t_nf_n$ is integrable, preserving convergence.
Also, the integrability condition (3.8) is reformulated using the distribution function:

$$\gamma(a) = \mu\{x \in S : |f(x)| \geq a\}.$$

Then, (3.8) holds if and only if $\gamma(a) < \infty$ for all $a > 0$, and

$$\int_0^1 \gamma(a)\ da < \infty.$$

Interestingly, the integral in condition (3.5) can remain finite even when the integral in (3.4) diverges. In fact, there are cases where the stricter requirement (3.8) fails entirely, so the sum Λ fails to converge absolutely but the right hand side of (3.5) remains finite. This implies that the sum Λ can be assigned a broader interpretation, allowing it to represent a well-defined random variable with finite variance despite the lack of absolute convergence. □

3.3 The characteristic functional

Begin by considering a special case of Campbell's Theorem, applying it to a non-negative measurable function $f \geq 0$ and choosing a particular value $\theta = -1$. Under this setting, equation (3.9) simplifies to:

$$\mathbb{E}(e^{-\Lambda_f}) = \exp\left\{-\int_S (1 - e^{-f(x)})d\mu(x)\right\}, \quad (3.12)$$

where:

$$\Lambda_f = \sum_{X \in \Pi} f(X).$$

The left-hand side of (3.12), viewed as a functional of the arbitrary function f, is known as the *characteristic functional* of the random process Π.
If equation (3.12) holds for a certain rich class $\mathcal{F}$ of functions, then it characterizes the process Π as a Poisson process with intensity measure μ. By rich, we mean that the class $\mathcal{F}$ is large enough to approximate or distinguish between different behaviors of the process. In particular, if it includes all simple functions, that is sufficient.

To make this concrete, suppose that for any function $f : S \to [0, \infty)$, we can define $\Lambda_f \leq \infty$. Now, assume f takes only finitely many values $f_1, f_2, \ldots, f_k$. That is, the space S can be partitioned into regions $A_j = \{x \in S : f(x) = f_j\}$. The intensity of each region is $m_j = \mu(A_j)$. Then, the expression for the characteristic functional becomes:

$$\mathbb{E}(e^{-\Lambda_f}) = \exp\left(-\sum_{j=1}^{k}(1 - e^{-f_j})m_j\right),$$

and

$$\Lambda_j = \sum_{j=1}^{k} f_j N(A_j).$$

Letting $z_j = e^{-f_j}$, this gives:

$$\mathbb{E}(z_1^{N(A_1)} z_2^{N(A_2)} \cdots z_k^{N(A_k)}) = \prod_{j=1}^{k} e^{m_j(z_j - 1)}.$$

Since this holds for $0 < z_j < 1$, the $N(A_j)$ are independent, with Poisson distributions $\mathcal{P}(m_j)$. Thus, Π is a Poisson process with mean measure μ.

To determine whether a random countable set is a Poisson process, it is sufficient to verify that equation (3.12) holds for a broad enough class of functions. This approach simplifies the task by focusing on a subset of functions rather than all possible functions.
A significant result regarding the characterization of the homogeneous Poisson process in one dimension was established by Alfréd Rényi in 1967. He proved that if a random subset of $\mathbb{R}$ has the property that, for any set A composed of a finite union of intervals, the number of points $N(A)$ follows a Poisson distribution with mean $\lambda|A|$, then the random set must be a Poisson process with constant rate λ.

Unlike the standard definition of a Poisson process, which requires independence of counts in disjoint intervals, Rényi's theorem shows that this property emerges automatically from the Poisson distribution assumption. Because the counts follow a Poisson process, they are automatically independent in non-overlapping regions. However, this independence is not guaranteed if one only assumes the Poisson property on individual intervals. A counterexample by Moran in 1967 illustrates this point: he constructed a non-Poisson process in which $N(A)$ follows a distribution $\mathcal{P}(\lambda|A|)$ for every interval A. This highlights the necessity of ensuring the Poisson distribution applies not just to single intervals but also to their finite unions.

Rényi's result aligns with the characteristic functional approach, though it doesn't directly provide a proof. Let's see why. For a homogeneous Poisson process on $\mathbb{R}$, the characteristic functional simplifies to:

$$\mathbb{E}(e^{\Lambda_f}) = \exp\left(-\lambda \int_{-\infty}^{\infty} (1 - e^{-f(x)})dx\right). \tag{3.13}$$

This identity must hold for all positive simple functions f.

Assuming that $N(A)$ follows a Poisson distribution with mean $\lambda|A|$ for all finite unions of intervals is the same as requiring that equation (3.13) holds for all positive simple functions that take only two values. Rényi's key idea was that if (3.13) holds for these two-valued simple functions, then it also holds for all positive simple functions. This gives a powerful shortcut for describing Poisson processes using minimal information.

3.4 Rényi's Theorem

Rényi's Theorem initially characterized the homogeneous Poisson process in one dimension, but it can be extended in multiple ways. For instance, the mean measure is not restricted to be a multiple of Lebesgue measure, and the state space can be more complex than $\mathbb{R}$. We will demonstrate its application to a random subset of $\mathbb{R}^d$ that remains finite within bounded regions.

A deeper exploration of this idea emerges from Kendall's (1974) work on random sets. His research highlights the significance of the *avoidance function*, defined as:

$$\alpha(A) = \mathbb{P}\{N(A) = 0\} = \mathbb{P}\{\Pi \cap A = \emptyset\},$$

where $\alpha(A)$ represents the probability that no points of the process Π fall within the set A.

For a Poisson process with mean measure μ, this function takes the form:

$$\alpha(A) = e^{-\mu(A)}. \tag{3.14}$$

Interestingly, it is not essential to assume that $N(A)$ follows a Poisson distribution. It's enough to assume that the probability of zero points in a set A has the form (3.14) for some measure μ.

This leads to a generalized version of Rényi's Theorem, framed in terms of *rectangles* in $\mathbb{R}^d$. Here, a rectangle is defined as a set of the form:

$$(a, b] = \{x = (x_1, \ldots, x_d) : a_i < x_i \leq b_i \text{ for all } i = 1, 2, \ldots, d\},$$

where $a, b \in \mathbb{R}^d$ and $a_i < b_i$ for each i.

Theorem 3.4 (**Rényi's Theorem**). *Let μ be a non-atomic measure on $\mathbb{R}^d$ that is finite on bounded sets. Consider a random countable subset Π of $\mathbb{R}^d$ with the property that for any finite union A of rectangles,*

$$\mathbb{P}(\Pi \cap A = \emptyset) = e^{-\mu(A)}. \tag{3.15}$$

Then Π is a Poisson process with mean measure μ.

This remarkable theorem offers two distinct approaches to establishing Poisson process structure, revealing the deep connection between Poisson distributions and independence.

In the first approach, begin with a random countable set Π, where for every finite union of rectangles A, the count $N(A)$ follows a Poisson distribution with finite mean. No independence is assumed at this stage. Then, $\mu(A) = \mathbb{E}(N(A))$ is a non-atomic measure on $\mathbb{R}^d$, finite on bounded sets. Since $N(A)$ is Poisson distributed, (3.15) holds, and Rényi's Theorem confirms that Π is a Poisson process.

For the second approach, start with a random countable set Π whose avoidance function satisfies $\alpha(A)$ for bounded A. For disjoint finite unions of rectangles A and B, the events

$$\{\Pi \cap A = \emptyset\} \text{ and } \{\Pi \cap B = \emptyset\}$$

are independent. Then

$$\alpha(A \cup B) = \mathbb{P}(\Pi \cap A = \Pi \cap B = \emptyset) = \alpha(A)\alpha(B)$$

implies that $\mu(A) = -\log\ \alpha(A)$ is finitely additive. If μ can be shown to be countably additive and non-atomic, the theorem applies, and Π is a Poisson process. To ensure μ is a non-atomic measure, it suffices that for any decreasing sequence of rectangles (A_n) shrinking to a single point or the empty set, $\lim \mu(A_n) = 0$. Hence, the only further condition we need to impose is that, for any sequence, $\lim \alpha(A_n) = 1$. This excludes sets Π with fixed points (i.e, points x where $\mathbb{P}(x \in \Pi) > 0$), which would violate the Poisson process definition.
Consider a trivial counterexample: let Π be a random set that is either $\emptyset$ (with probability 1/2) or $\{0\}$ (with probability 1/2). Then the avoidance function is:

$$\sigma(A) = \begin{cases} 1 & \text{if } 0 \notin A, \\ \frac{1}{2} & \text{if } 0 \in A. \end{cases}$$

While this example satisfies the positivity condition and the multiplicative property for disjoint sets, it fails the crucial limiting condition $\lim \alpha(A_n) = 1$. This occurs precisely because Π has a fixed point at the origin with probability $\frac{1}{2}$, then $\alpha(A_n) \not\to 1$ for rectangles A_n shrinking to $\{0\}$. Clearly, Π is not a Poisson process, as Poisson processes must be completely random without deterministic points.

The full proof of Rényi's Theorem involves technical steps beyond the scope of this discussion. While the theorem's conditions and implications are conceptually clear, the rigorous verification requires advanced tools from stochastic process theory.

Appendix A

Convergence Theorems

To support the proof of Campbell's Theorem (Theorem 3.1) in Section 3.2, we present two auxiliary results that are used in the argument. These theorems are stated without proof, as they are standard in the literature. [7] [10]

Theorem A.1 (**Dominated Convergence**). *Let $\{f_n\}_n$ be a sequence of measurable functions on a set S with respect to a measure μ, such that*

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \textit{for almost every } x \in S \textit{ (with respect to } \mu).$$

If there exists an integrable function g such that

$$|f_n(x)| \leq g(x) \quad \textit{for all } n \geq 1 \textit{ and almost every } x \in S,$$

and $\int_S g\, d\mu < \infty$, then:

$$\lim_{n\to\infty} \int_S f_n\, d\mu = \int_S \lim_{n\to\infty} f_n\, d\mu = \int_S f\, d\mu.$$

Theorem A.2 (**Monotone Convergence**). *Let $\{f_n\}_n$ be a sequence of measurable functions on a set S with respect to a measure μ, such that*

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \textit{for almost every } x \in S \textit{ (with respect to } \mu).$$

If

$$0 \leq f_1(x) \leq f_2(x) \leq \cdots \leq f_n(x) \leq \cdots \quad \textit{for almost every } x \in S,$$

then:

$$\lim_{n\to\infty} \int_S f_n\, d\mu = \int_S \lim_{n\to\infty} f_n\, d\mu = \int_S f\, d\mu.$$

Conclusions

This thesis explored the math behind Poisson processes—powerful models used to describe random events happening over time or space. We saw that the Poisson distribution plays two important roles: it accurately describes random, separate events (like phone calls coming into a call center), and it also shows up when we look at what happens to binomial models in rare-event situations. This makes it a natural and flexible tool in probability. The Poisson process is initially defined on $\mathbb{R}$, but it can be extended to more general spaces, as explored in this project.

Three main theorems: Superposition, Mapping, and Existence, showed how strong and adaptable the Poisson process is when we combine or transform it. These ideas aren't just useful for theory, they help in real-world problems too, like modeling internet traffic or lines in a store, where Poisson assumptions make things easier to analyze.

Campbell's and Rényi's theorems gave us the math tools to study things like averages and totals of Poisson events. They connect abstract math (like moments and characteristic functions) to real, measurable data. This is especially useful in areas like astrophysics (for detecting particles) or finance (for modeling sudden changes in prices).

In short, this thesis shows that the Poisson process is a powerful and flexible way to describe systems where independent, rare events happen, making it a kind of common language for many different scientific and practical problems.

Bibliography

[1] Pierre Brémaud and François Baccelli. *Elements of queueing theory: Palm Martingale calculus and stochastic recurrences*. Springer Science and Business Media, 2002. ISBN: 978-3-642-08537-6.

[2] CFI Team. *Poisson distribution*. Corporative Finance Institute. Aug. 2023. URL: https://corporatefinanceinstitute.com/resources/data-science/poisson-distribution/ (visited on 06/10/2025).

[3] H. Paul Keeler. *Notes on the Poisson point process*. Tech. rep. Tech. Rep. Berlin, Germany: Weierstrass Inst., 2016. URL: https://hpaulkeeler.com/wp-content/uploads/2018/08/PoissonPointProcess.pdf.

[4] John Frank Charles Kingman. *Poisson point processes*. Reprinted 2002. 104 p. p. 1-35. ISBN: 0-19-853693 3. New York: Oxford University Press, 1993.

[5] Günter Last and Mathew Penrose. *Lectures on the Poisson Process*. Cambridge University Press, 2017, pp. 1–24. ISBN: 978-1107088016.

[6] Penn State University. *Lesson 11: Poisson Random Variables. STAT 504: Categorical Data Analysis*. 2025. URL: https://online.stat.psu.edu/stat504/book/export/html/667.

[7] R. Timoney. *MA2224: Probability – Chapter 4 Notes*. 2023. URL: https://www.maths.tcd.ie/~richardt/MA2224/MA2224-ch4.pdf (visited on 06/10/2025).

[8] Wikipedia Contributors. *Campbell's theorem (probability)*. 2023. URL: https://en.wikipedia.org/wiki/Campbell%27s_theorem_(probability) (visited on 06/10/2025).

[9] Wikipedia Contributors. *Poisson point process*. May 2025. URL: https://en.wikipedia.org/wiki/Poisson_point_process (visited on 06/10/2025).

[10] Amie Wilkinson. *Monotone Convergence Theorem*. Nov. 2017. URL: https://www.math3ma.com/blog/monotone-convergence-theorem (visited on 06/10/2025).