5-2013

# A Study of Poisson and Related Processes with Applications

Phillip Mingola
*University of Tennessee - Knoxville*, pmingola@utk.edu

University of Tennessee, Knoxville

Honors Thesis

# A Study of Poisson and Related Processes with Applications

*Author:*
Phillip Mingola

*Advisor:*
Dr. Balram Rajput

*A thesis submitted in fulfilment of the Honors Thesis requirement in the Mathematics Department*

April 26, 2013

## 1. Introduction

The purpose of this thesis is to make a study with a view to better understand the Poisson and related processes (e.g., the compound Poisson process) and applications of these stochastic processes. I have compiled here some results that I believe would be useful and accessible to undergraduate math students. After researching the literature on the Poisson processes and related topics (e.g., binomial and Poisson random variables), I collected some relevant definitions and theorems, which are presented here. Because the Poisson process has many applications and constitutes a very wide subfield of stochastic processes, these notes are, by no means, meant to be comprehensive. Nevertheless, I feel the results contained herein should be useful to undergraduate math students who are interested in learning about Poisson processes and their applications.

The material is arranged as follows: Section 2 contains preliminary information about binomial and Poisson random variables, and a result of analysis; Section 3 contains properties of Poisson processes; Section 4 contains some applications of Poisson processes; Section 5 contains an introduction to the Compound Poisson process; Section 6 is the conclusion; and Section 7 lists references used.

Finally, I should note that the material contained in this thesis is standard, and I have borrowed the definitions and results from several sources [2, 3, 5, 7]. However, whenever I found that some part of a proof was difficult to understand, I searched the literature for clarification and included additional results with proofs; this addition makes the proofs presented here more complete, accessible, and clear at the undergraduate level.

## 2. Preliminaries

2.1. **Bernoulli Trials and Binomial Random Variables.** Consider a random experiment in which there are two possible outcomes; call one of these outcomes a "success" denoted by $S$ and the other a "failure" denoted by $F$. Suppose that the probability of success is $p \in (0,1)$, and so the probability of failure is $q = 1 - p$. This is called a **Bernoulli trial** with probability of $S$ equal to $p$.

Now, we define a random variable associated with a Bernoulli trial. Let $\mathscr{S} = \{S, F\}$ and $X : \mathscr{S} \to \{0, 1\}$ be a function on the sample space $\mathscr{S}$: $X(S) = 1$ and $X(F) = 0$. Then

$$\begin{aligned} \mathbf{p}_X(0) &= \mathbb{P}(X = 0) = q = 1 - p \\ \mathbf{p}_X(1) &= \mathbb{P}(X = 1) = p. \end{aligned}$$

This random variable $X$ is called a **Bernoulli random variable** with probability of $S = p$; $\mathbf{p}_X$ is its probability mass function.

If we perform $n$ independent, identical Bernoulli trials each with probability of $S$ equal to $p$, and let $Y$ be the number of successes. Then for $0 \leq k \leq n$, we have,

$$\mathbf{p}_Y(k) = P(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} = \binom{n}{k} p^k q^{n-k}. \tag{1}$$

To intuitively understand why this formula holds, suppose we record the value of $n$ independent Bernoulli trials. Since each trial is independent, the probability of $k$ successes is $p^k$ and the probability of $n-k$ failures is $q^{n-k}$. So the probability of $k$ successes followed by $n-k$ failures is given by $p^k q^{n-k}$. However, we do not only wish to consider cases where the first $k$ trials are successes followed by $n-k$ failures. Rather, we need to consider any sequence of successes and failures with exactly $k$ successes. The number of such sequences is $\binom{n}{k}$ and the probability of each of these sequences occurring is $p^k q^{n-k}$, so the probability of $k$ successes is given by (1). The random variable $Y$ with a probability mass function given by (1) is called a **binomial random variable** with parameters $n$ and $p$.

We can check that (1) does, in fact, define a probability mass function: It is clear that $\mathbf{p}_Y(x) = \mathbb{P}(Y = x) \geq 0$ for all $x \in \mathbb{R}$. Now, using the binomial theorem, $\displaystyle\sum_{k=0}^{n} \mathbf{p}_Y(k) = (p + (1-p))^n = 1$. Thus, $\mathbf{p}_Y$ defines a probability mass function. It is easy to show that the expected value of the binomial random variable $Y$ is $np$ and its variance is $npq$.

2.2. **Poisson Random Variables.** Computing probabilities for a binomial random variable $X$ is relatively easy for small values of $n$, but for large values of $n$, $\binom{n}{k} = \frac{n!}{(n-k)!k!}$ is cumbersome to compute and use. An important method to approximate $\mathbb{P}(X = k)$ for large $n$ is provided by what is called the Poisson approximation. We point out this approximation here: Let $X_n$ be a binomial random variable with parameters $n$ and $p_n = \frac{\lambda}{n}$ (where $\lambda > 0$ is given), then

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \frac{n!}{k!(n-k)!} p_n^k q_n^{n-k}, \qquad (q_n = 1 - p_n) \\
&= \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)(n-2)\cdots(n-k+2)(n-k+1)}{k!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{n(n-1)(n-2)\cdots(n-k+2)(n-k+1)}{n^k} \frac{\lambda^k}{k!} \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}.
\end{aligned}$$

Now, taking the limit of $\mathbb{P}(X_n = k)$ as $n \to \infty$, we find

$$\mathbb{P}(X_n = k) \ \to \ 1 \frac{\lambda^k}{k!} \frac{e^{-\lambda}}{1} = e^{-\lambda} \frac{\lambda^k}{k!}. \tag{2}$$

Thus, if $n$ is large and $p = \lambda/n$ is sufficiently small, then this formula can be used to calculate approximate values of the binomial probabilities. Poisson was primarily interested in approximating (1) for large values of $n$ and small values of $p$.

After Poisson published formula (2) as an approximation for the binomial probabilities, he died before L.V. Bortkiewicz showed that

$$\mathbf{p}(k) = \frac{e^{-\lambda}\lambda^k}{k!}, \quad \lambda > 0, k = 0, 1, 2, \ldots \tag{3}$$

defines a probability mass function of a random variable with parameter $\lambda$. In fact, one notes $\sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}(e^{\lambda}) = 1.$

A random variable whose probability mass function is given by (3) is called a **Poisson random variable** with parameter $\lambda$. Poisson random variables arise in application for events that occur in the long term, but where the probability of the event occurring in a short time is very small. This is called the law of rare events. Poisson random variables are useful for modeling these kinds of random rare phenomena; examples include the times at which a radioactive particle decays, customers arrive at a store, a machine part breaks and needs to be reset, or that accidents occur at a given intersection.

Now, we compute the expected value $\mathbb{E}[X]$ of a Poisson random variable with parameter $\lambda$:

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} ke^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k k}{k!} = e^{-\lambda}\lambda\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!} = e^{-\lambda}\lambda\sum_{j=0}^{\infty}\frac{\lambda^j}{j!} = \lambda e^{-\lambda}e^{\lambda} = \lambda.$$

Similarly, we find that $\mathbb{E}[X^2] = \lambda^2 + \lambda$, so the variance of $X$, $\text{var}[X]$, is $\mathbb{E}[X^2] - \mathbb{E}^2[X] = \lambda^2 + \lambda - \lambda^2 = \lambda$.

The following theorem shows an interesting property of Poisson random variables.

**Theorem 1.** *Let $X, Y$ be independent Poisson random variables with parameters $\lambda, \mu$. Then $X + Y$ is a Poisson random variable with parameter $\lambda + \mu$.*

*Proof.* Consider $\mathbb{P}(X+Y = k) = \mathbb{P}\left(\bigcup_{j=0}^{k}\{X = j, Y = k - j\}\right)$. Because the events $\{X = j, Y = k - j\}_{0\leq j\leq k}$ are mutually exclusive and $X$ and $Y$ are independent, we have

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

□

The following theorem shows a relationship between Poisson random variables and Binomial random variables:

**Theorem 2.** *Let $N$ be a Poisson random variable with parameter $\lambda$, and let $M$ be a binomial random variable conditioned on $N$ with parameters $N, p$. Then the distribution of $M$ is Poisson with parameter $\lambda p$.*

*Proof.* By the law of total probability, $\mathbb{P}(M = k) = \sum_{n=0}^{\infty} \mathbb{P}(M = k|N = n)\mathbb{P}(N = n)$. For $n < k$, this probability is zero, so changing the index of our summation and plugging in the probability mass functions, we have $\mathbb{P}(M = k) = \sum_{n=k}^{\infty} \left[\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\right]\left(\frac{\lambda^n e^{-\lambda}}{n!}\right) =$

$$\frac{e^{-\lambda}\lambda^k p^k}{k!}\sum_{n=k}^{\infty}\left(\frac{n!(1-p)^{n-k}}{(n-k)!}\frac{\lambda^{n-k}}{n!}\right) = \frac{e^{-\lambda}(\lambda p)^k}{k!}\sum_{n=0}^{\infty}\frac{((1-p)\lambda)^n}{n!} = \frac{e^{-\lambda}(\lambda p)^k}{k!}e^{\lambda-\lambda p} = \frac{e^{-\lambda p}(\lambda p)^k}{k!},$$

as desired. □

2.3. **A Formula for Computation of Probabilities.**

**Proposition 1.** *Let $f$ be the joint continuous probability density function of the random variables $X_1, X_2, \ldots, X_n$, then for any $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ and $\Delta x_1, \Delta x_2, \ldots, \Delta x_n$, with $\Delta x = (\Delta x_1) \cdots (\Delta x_n)$, we have*

$$\mathbb{P}\left(\bigcap_{i=1}^{n}(x_i \leq X_i \leq x_i + \Delta x_i)\right) = f(x)\Delta x + o(\Delta x),$$

*as $\Delta x_i \to 0$ for $i = 1, 2, \ldots, n$.*

*Proof.* By the definition of the probability density function, we can compute the left hand side using integrals:

$$(4)\ \mathbb{P}\left(\bigcap_{i=1}^{n}(x_i \leq X_i \leq x_i + \Delta x_i)\right) = \int_{x_n}^{x_n+\Delta x_n} \cdots \int_{x_1}^{x_1+\Delta x_1} f(y_1, y_2, \ldots, y_n)dy_1dy_2\ldots dy_n.$$

By [1, p. 299] there is some $\mathbf{x}' = (x_1', x_2', \ldots, x_n')$ such that $x_i \leq x_i' \leq x_i + \Delta x_i$ and such that the right hand side of (4) is equal to $f(\mathbf{x}')\Delta x$. Thus, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^{n}(x_i \leq X_i \leq x_i + \Delta x_i)\right) &= f(\mathbf{x}')\Delta x \\
&= (f(\mathbf{x}') + f(\mathbf{x}) - f(\mathbf{x}))\Delta x \\
&= f(\mathbf{x})\Delta x + (f(\mathbf{x}') - f(\mathbf{x}))\Delta x \\
&= f(\mathbf{x})\Delta x + o(\Delta x).
\end{aligned}$$

The final equality holds because as $\Delta x_i \to 0$ for all $i$, we have that $\mathbf{x}' \to \mathbf{x}$, and by continuity of $f$. $f(\mathbf{x}') - f(\mathbf{x}) = \dfrac{(f(\mathbf{x}') - f(\mathbf{x}))\Delta x}{\Delta x} \to 0$.

**Remark**: The above result, in fact, holds when $f$ is Lebesgue measurable. Precisely, in this case, the conclusion of Proposition 1 holds for all $\mathbf{x} \in A \subset \mathbb{R}^n$ with $\mathscr{L}(A^c) = 0$, where $A$ is a Borel set. The proof is similar to the above; one uses Theorem 8.6 of Rudin instead of [1, p. 299]. $\square$

## 3. The Poisson Process

3.1. **Definition of the Poisson Process.** In general, a random phenomenon that evolves in time is called a stochastic process. More precisely, a stochastic process is a family of random variables $\{X_t : t \in T\}$, with some associated indexing set $T$. As we will be studying the Poisson Process in this thesis, we will primarily be interested in the case when $T = (0, \infty)$.

Consider a sequence of random rare events that occur at various times. Set some time $t > 0$, then we can define a counting function $N(t)$ that gives the number of events that occur in the time interval $(0, t]$. For any value of $t$, $N(t)$ is a discrete random variable. We shall make the following assumptions on $N(t)$ and on our random events. **Stationarity**: If $(a, b]$ and $(c, d]$ are disjoint time intervals where $b - a = d - c$ (i.e., the time intervals have the same length), then the probability of $n$ events occurring in the interval $(a, b]$ is equal to the probability of $n$ events occurring in $(c, d]$ for all $n \geq 0$. **Independent Increments**: For all $n \geq 0$, the probability of $n$ events occurring in the interval $(t, t + s]$, for $s > 0$, is independent of $N(t)$, the number of events occurring in $(0, t]$. **Orderliness**: For small time intervals, the probability of more than one event occurring is very small. That is to say $\lim_{h \to 0} \frac{\mathbb{P}(N(h) > 1)}{h} = 0$ or $\mathbb{P}(N(h) > 1) = o(h)$. The following theorem is very important, but its proof is lengthy and beyond the scope of this paper.

**Theorem 3.** *If random events occurring in time satisfy stationarity, independent increments, and orderliness, let* $N(t)$ *be the number of events that have occurred in* $(0, t]$*;* $N(0) = 0$*; and assume additionally* $0 < \mathbb{P}(N(t) = 0) < 1$ *for all* $t > 0$*, then there is some* $\lambda > 0$ *such that*

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

*Equivalently,* $N(t)$ *is a Poisson random variable with parameter* $\lambda t$*, and thus,* $\mathbb{E}[N(t)] = \lambda t$*, so* $\lambda = \mathbb{E}[N(1)]$.

The stochastic process $\{N(t) : t \geq 0\}$ defined in the above theorem is called a **Poisson process of intensity** $\lambda$.

3.2. **Properties of the Poisson Process.** The Poisson process has many useful properties, some of which are illustrated by the following theorems.

**Theorem 4.** *Let $N(t)$ be a Poisson process of intensity $\lambda$. Fix $t > 0$, and let $N(t) = n$. For fixed $u \in (0, t]$, the conditional distribution of $N(u)$ given that $N(t) = n$ is binomial with parameters $n, u/t$.*

*Proof.* Choose $k$ between 0 and $n$. Then we wish to compute $\mathbb{P}(N(u) = k | N(t) = n)$. We note that this is the same as $\dfrac{\mathbb{P}(N(u) = k, N(t) = n)}{\mathbb{P}(N(t) = n)}$. The probability $\mathbb{P}(N(u) = k, N(t) = n)$ is equal to $\mathbb{P}(N(u) = k, N(t) - N(u) = n - k)$. The event $\{N(t) - N(u) = n - k\}$ means that in the interval $(u, t]$, $n - k$ events occur, and by stationarity, $N(t) - N(u)$ has the same distribution as $N(t - u)$, and by independent increments, $N(t) - N(u)$ and $N(u)$ are independent. Thus, our equation simplifies to

$$\begin{aligned}
\mathbb{P}(N(u)=k|N(t)=n) &= \frac{\mathbb{P}(N(u)=k)\mathbb{P}(N(t-u)=n-k)}{\mathbb{P}(N(t)=n)} \\
&= \frac{\frac{(\lambda u)^k e^{-\lambda u}}{k!}\frac{(\lambda(t-u))^{n-k}e^{-\lambda(t-u)}}{(n-k)!}}{\frac{(\lambda t)^n e^{-\lambda t}}{n!}} \\
&= \frac{n!\lambda^k u^k \lambda^{n-k}(t-u)^{n-k}e^{-\lambda u}e^{-\lambda(t-u)}}{k!(n-k)!\lambda^n t^n e^{-\lambda t}} \\
&= \binom{n}{k}\frac{\lambda^n e^{-\lambda t}}{\lambda^n e^{-\lambda t}}\frac{u^k(t-u)^{n-k}}{t^k t^{n-k}} \\
&= \binom{n}{k}\left(\frac{u}{t}\right)^k\left(1-\frac{u}{t}\right)^{n-k}.
\end{aligned}$$

Thus, for given $N(t)=n$, $N(u)$ is a binomial random variable with parameters $n$, $\frac{u}{t}$. □

**Theorem 5.** *Suppose $N_1(t), N_2(t)$ are independent Poisson processes with intensities $\lambda, \mu$. The conditional distribution of $N_1(t)$ given $N_1(t)+N_2(t)=n$ is binomial with parameters $n$, $\frac{\lambda}{\lambda+\mu}$.*

*Proof.* Fix some $k$ between 0 and $n$. Then by independence of $N_1(t)$ and $N_2(t)$, we have

$$\mathbb{P}(N_1(t)=k|N_1(t)+N_2(t)=n) = \frac{\mathbb{P}(N_1(t)=k)\mathbb{P}(N_2(t)=n-k)}{\mathbb{P}(N_1(t)+N_2(t)=n)},$$

and by Theorem 1 (i.e., $N_1(t)+N_2(t)$ has a Poisson distribution with parameter $t(\lambda+\mu)$),

$$\begin{aligned}
\mathbb{P}(N_1(t)=k|N_1(t)+N_2(t)=n) &= \frac{\frac{(\lambda t)^k e^{-\lambda t}}{k!}\frac{(\mu t)^{n-k}e^{-\mu t}}{(n-k)!}}{\frac{(t(\mu+\lambda))^n e^{-t(\lambda+\mu)}}{n!}} \\
&= \frac{n!\lambda^k t^k \mu^{n-k} t^{n-k} e^{-t(\lambda+\mu)}}{k!(n-k)!t^n(\lambda+\mu)^k e^{-t(\lambda+\mu)}} \\
&= \binom{n}{k}\frac{\lambda^k}{(\lambda+\mu)^k}\frac{\mu^{n-k}}{(\lambda+\mu)^{n-k}} \\
&= \binom{n}{k}\left(\frac{\lambda}{\lambda+\mu}\right)^k\left(1-\frac{\lambda}{\lambda+\mu}\right)^{n-k}.
\end{aligned}$$

Hence, the conditional distribution of $N_1(t)$ given that $N_1(t)+N_2(t)=n$ is binomial with parameters $n$, $\frac{\lambda}{\lambda+\mu}$. □

In addition to being interested in the number of occurrences of a rare event, we will also want to study the times at which rare events occur. While the preceding results showed relationships between Poisson processes and binomial random variables, the following results show a connection between Poisson processes and the gamma and uniform distributions.

**Theorem 6.** *Consider a Poisson process $\{N(t) : t \geq 0\}$ with parameter $\lambda$, and let the time of the $i$th occurrence of an event be denoted by $S_i$. Then the probability density function of $S_n$ is*

$$f_{S_n}(t)=\frac{\lambda^n t^{n-1}}{(n-1)!}e^{-\lambda t}, \qquad n=1,2,\ldots,\ t\geq 0.$$

A particular case of this theorem is that for $n = 1$, we have that the time until the first occurrence has density $f_{S_1}(t) = \lambda e^{-\lambda t}$, which is an exponential random variable with parameter $\lambda$.

*Proof.* First we will compute the distribution function for $S_n$. The event $\{S_n \le t\}$ is the same as $\{N(t) \ge n\}$. So, $F_{S_n}(t) = \mathbb{P}(S_n \le t) = \mathbb{P}(N(t) \ge n) = \displaystyle\sum_{i=n}^{\infty} \frac{e^{-\lambda t}(\lambda t)^i}{i!} = 1 - \sum_{i=0}^{n-1} \frac{e^{-\lambda t}(\lambda t)^i}{i!}$. Now taking the derivative of $F_{S_n}$ with respect to $t$, we have

$$
\begin{aligned}
\frac{dF_{S_n}}{dt} = -\frac{d}{dt}\left[e^{-\lambda t}\sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!}\right] &= \lambda e^{-\lambda t}\sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!} - \lambda e^{-\lambda t}\sum_{i=0}^{n-1}\frac{i(\lambda t)^{i-1}}{i!} \\
&= \lambda e^{-\lambda t}\left[\sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!} - \sum_{i=1}^{n-1}\frac{(\lambda t)^{i-1}}{(i-1)!}\right] \\
&= \lambda e^{-\lambda t}\left[\sum_{i=0}^{n-1}\frac{(\lambda t)^i}{i!} - \sum_{i=0}^{n-2}\frac{(\lambda t)^i}{i!}\right] \\
&= \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} = \frac{\lambda^n t^{n-1}}{(n-1)!}e^{-\lambda t};
\end{aligned}
$$

this proves the result.

□

**Theorem 7.** *Consider the inter arrival times $I_1, I_2, \ldots, I_n$ with $I_k = S_k - S_{k-1}$ and $S_0 = 0$. These inter arrival times are independent each having an exponential distribution with parameter $\lambda$; thus,*

$$f_{I_k}(t) = \lambda e^{-\lambda t}.$$

*Proof.* We shall prove this only for the case of $n = 2$ since the proof in the general case is essentially the same. That the $I_k$'s are independent is equivalent to saying that their joint probability density function is the product of their individual probability density functions. So we need to show that $f_{I_1,I_2}(t_1,t_2) = (\lambda e^{\lambda t_1})(\lambda e^{\lambda t_2}) = \lambda^2 e^{\lambda(t_1+t_2)}$. The event that $t_1 < I_1 < t_1 + \Delta t_1$ and $t_2 < I_2 < t_2 + \Delta t_2$ is equivalent to the condition that $t_1 < S_1 < t_1 + \Delta t_1$ and $t_1 + \Delta t_1 + t_2 < S_2 < t_1 + \Delta t_1 + t_2 + \Delta t_2$. Using independence of increments, stationarity, and Proposition 1, we have:

$$
\begin{aligned}
f_{I_1,I_2}(t_1,t_2)\Delta t_1 \Delta t_2 &= \mathbb{P}(t_1 < S_1 < t_1 + \Delta t_1, t_1 + \Delta t_1 + t_2 < S_2 < t_1 + \Delta t_1 + t_2 + \Delta t_2) + o(\Delta t_1 \Delta t_2) \\
&= \mathbb{P}(N(t_1) = 0)\mathbb{P}(N(\Delta t_1) = 1)\mathbb{P}(N(t_2) = 0)\mathbb{P}(N(\Delta t_2) = 1) + o(\Delta t_1 \Delta t_2) \\
&= e^{\lambda t_1} e^{\lambda \Delta t_1}(\lambda \Delta t_1) e^{\lambda t_2} e^{\lambda \Delta t_2}(\lambda \Delta t_2) + o(\Delta t_1 \Delta t_2) \\
&= \lambda^2 e^{\lambda(t_1+t_2+\Delta t_1+\Delta t_2)} \Delta t_1 \Delta t_2 + o(\Delta t_1 \Delta t_2).
\end{aligned}
$$

Dividing this whole equation by $\Delta t_1 \Delta t_2$ and taking the limit as $\Delta t_1 \to 0$ and $\Delta t_2 \to 0$ yields the desired result. Note that we have implicitly assumed that $f$ is continuous, but this condition can be removed in light of the remark following Proposition 1. □

**Theorem 8.** *Let $\{N(t) : t \geq 0\}$ be a Poisson process with intensity $\lambda$, and suppose that for fixed $t$, $N(t) = n$. For $1 \leq i \leq n$, define $S_i$ to be the time of the occurrence of the $i$th event.*

*Then the conditional joint probability density function of* $S_1, \ldots, S_n$ *given that* $N(t) = n$ *is* $f_{S_1,\ldots,S_n|N(t)=n}(t_1, t_2, \ldots, t_n) = \frac{n!}{t^n}$*, with* $0 < t_1 < t_2 < \cdots < t_n < t$.

As the proof of the general case of this theorem is tedious, we will only prove it for the case of $n = 2$. The proof of the general case is similar.

*Proof.* Suppose for fixed $t$, $N(t) = n = 2$. Then choose $\Delta t_1 > 0$, $\Delta t_2 > 0$ so that $0 < t_1 < t_1 + \Delta t_1 < t_2 < t_2 + \Delta t_2 < t$. The event that $t_i < S_i \leq t_i + \Delta t_i$ for $i = 1, 2$, given $N(t) = 2$ means that one event occurs in $(t_1, t_1 + \Delta t_1]$, one event occurs in $(t_2, t_2 + \Delta t_2]$, and no events occur in $(0, t_1] \cup (t_1 + \Delta t_1, t_2] \cup (t_2 + \Delta t_2, t]$. The joint probability density function for $S_1, S_2$ conditional on $N(t) = 2$ satisfies $f_{S_1,S_2|N(t)=2}(t_1, t_2)\Delta t_1 \Delta t_2 = \mathbb{P}(t_1 < S_1 \leq t_1 + \Delta t_1, t_2 < S_2 \leq t_2 + \Delta t_2 | N(t) = 2) + o(\Delta t_1, \Delta t_2)$, by Proposition 1. Using independence of increments and stationarity, we have:

$$f_{S_1,S_2|N(t)=2}(t_1, t_2)\Delta t_1 \Delta t_2 = \frac{\mathbb{P}(t_1 < S_1 < t_1 + \Delta t_1, t_2 < S_2 < t_2 + \Delta t_2, N(t) = 2)}{\mathbb{P}(N(t) = 2)} + o(\Delta t_1, \Delta t_2)$$

$$= \frac{\mathbb{P}(N(t_1)=0)\mathbb{P}(N(\Delta t_1)=1)\mathbb{P}(N(t_2-(t_1+\Delta t_1))=0)\mathbb{P}(N(\Delta t_2)=1)P(N(t-(t_2+\Delta t_2))=0)}{\mathbb{P}(N(t)=2)}$$

$$+o(\Delta t_1, \Delta t_2)$$

$$\begin{aligned}
&= \frac{\frac{e^{-\lambda t_1}(\lambda t_1)^0}{0!}\frac{e^{-\lambda \Delta t_1}(\lambda \Delta t_1)^1}{1!}\frac{e^{-\lambda(t_2-t_1-\Delta t_1)}(\lambda(t_2-t_1-\Delta t_1)^0}{0!}\frac{e^{-\lambda \Delta t_2}(\lambda \Delta t_2)^1}{1!}\frac{e^{-\lambda(t-t_2-\Delta t_2)}(\lambda(t-t_2-\Delta t_2))^0}{0!}}{\frac{e^{-\lambda t}(\lambda t)^2}{2}} + o(\Delta t_1, \Delta t_2) \\
&= \frac{e^{-\lambda t_1}e^{-\lambda \Delta t_1}\lambda \Delta t_1 e^{-\lambda(t_2-t_1-\Delta t_1)}e^{-\lambda \Delta t_2}\lambda \Delta t_2 e^{-\lambda(t-t_2-\Delta t_2)}}{\frac{e^{-\lambda t}(\lambda t)^2}{2}} + o(\Delta t_1, \Delta t_2) \\
&= \frac{2e^{-\lambda(t_1+\Delta t_1+t_2-\Delta t_1-t_1+\Delta t_2+t-t_2-\Delta t_2)}\lambda^2 \Delta t_1 \Delta t_2}{e^{-\lambda t}\lambda^2 t^2} + o(\Delta t_1, \Delta t_2) = \frac{2e^{-\lambda t}\Delta t_1 \Delta t_2}{e^{-\lambda t}t^2} + o(\Delta t_1, \Delta t_2) \\
&= \frac{2!\Delta t_1 \Delta t_2}{t^2} + o(\Delta t_1, \Delta t_2).
\end{aligned}$$

Dividing through the entire equation by $\Delta t_1 \Delta t_2$ and taking the limit as $\Delta t_1 \to 0, \Delta t_2 \to 0$ gives the desired result. Again, note that we have assumed $f_{S_1,S_2|N(t)=n}$ to be continuous, but this condition can be removed in light of the remark after Proposition 1. □

Fix some $t > 0$ and consider $n$ independent, identically distributed uniform random variables on $(0, t]$: $U_1, U_2, \ldots, U_n$. Then we define $U_{(1)} = \min\{U_1, U_2, \ldots, U_n\}$, $U_{(2)} = \min\{\{U_1, U_2, \ldots, U_n\} \setminus \{U_{(1)}\}\}$, $U_{(3)} = \min\{\{U_1, U_2, \ldots, U_n\} \setminus \{U_{(1)}, U_{(2)}\}\}$, and so on. The random variables $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ are called the **order statistics** of $U_1, U_2, \ldots, U_n$.

Since it is well known that the joint probability density function of $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$ is $f_{U_{(1)},U_{(2)},\ldots,U_{(n)}}(x_1, x_2, \ldots, x_n) = \frac{n!}{t^n}$ for $0 < x_1 < x_2 < \cdots < x_n \le t$ and equal to 0 otherwise. Therefore, it follows from Theorem 8 that the conditional distribution of the arrival times

$S_1, S_2, \dots, S_n$ given that $N(t) = n$ is the same as the distribution of the order statistics of $U_1, U_2, \dots, U_n$ (i.e, that of $U_{(1)}, U_{(2)}, \dots, U_{(n)}$). Suppose now that $N(t) = n$. Let $Y_1, Y_2, \dots, Y_n$ be independent and identically distributed according to $G(y) = \mathbb{P}(Y_k \leq y)$, and suppose also that $Y_1, Y_2, \dots, Y_n$ are independent of $S_1, S_2, \dots, S_n$; $U_1, U_2, \dots, U_n$; and $N(t)$. Define a new random variable $I_k(t) = 1$ if and only if $S_k + Y_k \geq t$, and $I_k(t) = 0$ otherwise. Similarly, we define $J_k(t) = 1$ if $U_{(k)} + Y_k \geq 1$ and $J_k(t) = 0$ otherwise, and $L_{j,k}(t) = 1$ if $U_j + Y_k \geq 1$ and $L_{j,k}(t) = 0$ otherwise (if $j = k$, we will write $L_{k,k} = L_k$). In the following, we shall use the notations introduced above.

**Proposition 2.** *The probability mass function of* $\sum_{k=1}^{n} L_k(t)$ *is equal to the conditional probability mass function of* $\sum_{k=1}^{n} I_k(t)$ *given* $N(t) = n$.

*Proof.* Let $\Sigma_n$ be the set of all permutations of $(1, 2, \dots, n)$. Let $\sigma = (\sigma(1), \sigma(2), \dots, \sigma(n)) \in \Sigma_n$ and let $A_\sigma = \{U_{\sigma(1)} < U_{\sigma(2)} < \cdots < U_{\sigma(n)}\}$. We first make two observations:

i ) The joint distribution of $\{U_{(1)}, U_{(2)}, \dots, U_{(n)}, Y_1, Y_2, \dots, Y_n\}$ is the same as the conditional joint distribution of $\{S_1, S_2, \dots, S_n, Y_1, Y_2, \dots, Y_n\}$ given that $N(t) = n$. This follows as shown below: Let $s_1, s_2, \cdots, s_n \in \mathbb{R}$ and $y_1, y_2, \dots, y_n \in \mathbb{R}$. Then using the independence of $\{Y_1, \dots, Y_n\}$, $\{U_1, U_2, \dots, U_n\}$, and $N(t)$ and Theorem 8, we have

$\mathbb{P}(S_1 \leq s_1, S_2 \leq s_2, \ldots, S_n \leq s_n, Y_1 \leq y_1, \ldots, Y_n \leq y_n | N(t) = n)$

$$\begin{aligned}
&= \frac{\mathbb{P}(S_1 \leq s_1, \ldots, S_n \leq s_n, Y_1 \leq y_1, \ldots, Y_n \leq y_n, N(t) = n)}{\mathbb{P}(N(t) = n)} \\
&= \frac{\mathbb{P}(S_1 \leq s_1, \ldots, S_n \leq s_n, N(t) = n)\mathbb{P}(Y_1 \leq y_1, \ldots, Y_n \leq y_n)}{\mathbb{P}(N(t) = n)} \\
&= \mathbb{P}(S_1 \leq s_1, \ldots, S_n \leq s_n | N(t) = n)\mathbb{P}(Y_1 \leq y_1, \ldots, Y_n \leq y_n) \\
&= \mathbb{P}(U_{(1)} \leq s_1, \ldots, U_{(n)} \leq s_n)\mathbb{P}(Y_1 \leq y_1, \ldots, Y_n \leq y_n) \\
&= \mathbb{P}(U_{(1)} \leq s_1, \ldots, U_{(n)} \leq s_n, Y_1 \leq y_1, \ldots, Y_n \leq y_n).
\end{aligned}$$

ii ) The joint distribution of $\{(U_{\sigma(i)}, Y_i), i = 1, \ldots, n\}$ is the same as that of $\{(U_{\sigma(i)}, Y_{\sigma(i)}), i = 1, \ldots, n\}$: Let $s_1, s_2, \ldots, s_n \in \mathbb{R}$ and $y_1, y_2, \ldots, y_n \in \mathbb{R}$, then because $Y_1, Y_2, \ldots, Y_n$ are identically distributed and independent of the $U_i$'s, we have

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i=1}^{n} \left(U_{\sigma(i)} \leq s_i, Y_i \leq y_i\right)\right) &= \mathbb{P}\left(\left(\bigcap_{i=1}^{n} U_{\sigma(i)} \leq s_i\right) \cap \left(\bigcap_{i=1}^{n} Y_i \leq y_i\right)\right) \\
&= \mathbb{P}\left(\bigcap_{i=1}^{n} U_{\sigma(i)} \leq s_i\right) \mathbb{P}\left(\bigcap_{i=1}^{n} Y_i \leq y_i\right) \\
&= \mathbb{P}\left(\bigcap_{i=1}^{n} U_{\sigma(i)} \leq s_i\right) \mathbb{P}\left(\bigcap_{i=1}^{n} Y_{\sigma(i)} \leq y_i\right) \\
&= \mathbb{P}\left(\left(\bigcap_{i=1}^{n} U_{\sigma(i)} \leq s_i\right) \cap \left(\bigcap_{i=1}^{n} Y_{\sigma(i)} \leq y_i\right)\right) \\
&= \mathbb{P}\left(\bigcap_{i=1}^{n} \left(U_{\sigma(i)} \leq s_i, Y_{\sigma(i)} \leq y_i\right)\right).
\end{aligned}$$

Now, note that by Theorem 8 and (i),

$$\mathbb{P}\left(\sum_{k=1}^{N(t)} I_k(t) = m | N(t) = n\right) = \mathbb{P}\left(\sum_{k=1}^{n} J_k(t) = m\right).$$

Finally,

$$\begin{aligned}
\mathbb{P}\left(\sum_{k=1}^{n} J_k(t) = m\right) &= \sum_{\sigma\in\Sigma_n} \mathbb{P}\left(\sum_{k=1}^{n} J_k(t) = m | A_\sigma\right) \mathbb{P}(A_\sigma) \\
&= \sum_{\sigma\in\Sigma_n} \mathbb{P}\left(\sum_{k=1}^{n} L_{\sigma(k),k}(t) = m\right) \mathbb{P}(A_\sigma) \\
&= \sum_{\sigma\in\Sigma_n} \mathbb{P}\left(\sum_{k=1}^{n} L_{\sigma(k),\sigma(k)}(t) = m\right) \mathbb{P}(A_\sigma) \quad \text{(by (ii))} \\
&= \sum_{\sigma\in\Sigma_n} \mathbb{P}\left(\sum_{k=1}^{n} L_k(t) = m\right) \mathbb{P}(A_\sigma) \quad \text{(by symmetry)} \\
&= \mathbb{P}\left(\sum_{k=1}^{n} L_k(t) = m\right) \sum_{\sigma\in\Sigma_n} \mathbb{P}(A_\sigma) \\
&= \mathbb{P}\left(\sum_{k=1}^{n} L_k(t) = m\right).
\end{aligned}$$

This completes our proof.

□

## 4. Applications of the Poisson Process

The Poisson process is useful in modeling a large number of random phenomena. Several of these applications are presented here. Much of the material in this section comes from [3].

4.1. **Radioactive Decay.** One useful application of the Poisson process is in modeling the number of radioactive particles produced by a sample. Given some large collection of radioactive nuclei that emit $\alpha$-particles in time and for times much less than the half life of the radioactive substance (this can range from a fraction of a second to billions of years), the number of $\alpha$-particles emitted up to time $t$ can be modeled by a Poisson process, $\{N(t) : t \geq 0\}$ with intensity $\lambda$. Suppose the lifetimes of $\alpha$-particles are $Y_1, Y_2, \ldots$, and that these lifetimes are independent and identically distributed according to $G(y) = \mathbb{P}(Y_k \leq y)$. Let $S_k$ be the time that the $k$th particle is produced by the sample. Fix a time $t$, and suppose that $N(t) = n$. Then for $k \leq n$, we have that $S_k \leq t$. The $k$th $\alpha$-particle still exists (it has not yet decomposed) if and only if $S_k + Y_k \geq t$.

Recall the definitions of $I_k, J_k, L_{j,k}$ from the end of Section 3. Let $M(t) = \sum_{k=1}^{n} I_k(t)$ be the number of $\alpha$-particles still in existence at time $t$. Using symmetry of the particles and Proposition (2), we have

$$\mathbb{P}\left(M(t) = m | N(t) = n\right) = \mathbb{P}\left(\sum_{k=1}^{N(t)} I_k(t) = m | N(t) = n\right) = \mathbb{P}\left(\sum_{k=1}^{n} L_k(t) = m\right).$$

Clearly, the random variable $\sum_{k=1}^{n} L_k(t)$ has a binomial distribution with parameters $n$ and $p$, where

$$\begin{aligned}
p &= \mathbb{P}(U_k + Y_k \geq t) = \int_0^t \mathbb{P}(Y_k \geq t-u | U_k = u) f_{U_k}(u)\, du = \frac{1}{t}\int_0^t \mathbb{P}(Y_k \geq t-u) du \\
&= \frac{1}{t}\int_0^t [1 - \mathbb{P}(Y_k \leq t-u)] du = \frac{1}{t}\int_0^t [1 - G(t-u)] du \\
&= \frac{1}{t}\int_t^0 [1 - G(z)](-dz) = \frac{1}{t}\int_0^t [1-G(z)] dz.
\end{aligned}$$

The value of $p$ is easy to compute given an expression for $G(y)$. Now, we can easily compute the probability that there are $m$ $\alpha$-particles still in existence at time $t$:

$$\begin{aligned}
\mathbb{P}(M(t) = m) &= \sum_{n=m}^{\infty} \mathbb{P}(M(t) = m | N(t) = n) \mathbb{P}(N(t) = n) \\
&= \sum_{n=m}^{\infty} \binom{n}{m} p^m (1-p)^{n-m} \frac{(\lambda t)^n e^{-\lambda t}}{n!} = \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \frac{(\lambda t)^n e^{-\lambda t}}{n!} \\
&= \frac{e^{-\lambda t}}{m!} (p\lambda t)^m \sum_{n=m}^{\infty} \frac{(1-p)^{n-m} (\lambda t)^{n-m}}{(n-m)!} = \frac{e^{-\lambda t}}{m!} (p\lambda t)^m \sum_{n=0}^{\infty} \frac{(1-p)^j (\lambda t)^j}{j!} \\
&= \frac{e^{-\lambda t}}{m!} (p\lambda t)^m \sum_{n=0}^{\infty} \frac{[(1-p)(\lambda t)]^j}{j!} = \frac{e^{-\lambda t}}{m!} (p\lambda t)^m e^{(1-p)\lambda t} = \frac{e^{-p\lambda t} (p\lambda t)^m}{m!}.
\end{aligned}$$

Thus $M(t)$ is a Poisson random variable with parameter $(\lambda p)t$. Using our expression for $p$, we have $\lambda p t = \lambda \int_0^t [1 - G(y)] dy$. For a nonnegative continuous random variable $Y$, $\mathbb{E}[Y] = \int_0^\infty [1 - F_Y(y)] dy$, where $F_Y(y)$ is the distribution function for $Y$. Thus, as $t \to \infty$, $\int_0^t [1 - G(y)] dy \to \mathbb{E}[Y_k]$. Therefore, if we let $\mu = \mathbb{E}[Y_k] = \mathbb{E}[Y_1]$, it follows that for very large $t$, $M(t)$ has a Poisson distribution with parameter $\mu\lambda$. An important part of this result, is that on a large enough time scale, $M(t)$ does not depend directly on $G(y)$, but rather only on $\mu$, the mean value of $Y_k$.

4.2. **The Coupon Collection Problem.** Consider this problem as presented by Ross [pg. 268-269]: Suppose there are $m$ different types of coupons, and for a given coupon, the probability of the coupon being of type $j$ is $p_j$. Because there are only $m$ different types of coupons, we have $\sum_{j=1}^{m} p_j = 1$. Further, assume that the occurrences of the coupons are independent; that is to say, the occurrence of one coupon type has no bearing on the occurrence of any other coupon type. We wish to determine the expected number of coupons $\mathbb{E}(M)$ that one must collect in order to collect all $m$ types of coupons.

Let $M_j$ denote the number of coupons one needs to collect in order to collect one coupon of type $j$. Then clearly, $M = \max\{M_j\}$. The random variables $M_j$ are not independent, so it is not immediately evident how we can compute $M$. Suppose the coupons arrive according to a Poisson process of intensity 1. The event of this Poisson process is of type $j$ if the coupon is of type $j$. We now define a family of independent Poisson processes $\{N_j(t) : t \geq 0\}_{1\leq j\leq m}$, where the intensity of the $j$th process is $p_j$. Define $S_j$ to be the time of the first arrival of the $j$th coupon. That is to say, $S_j$ is the smallest value such that $N_j(S_j) = 1$. The time we are interested is $S = \max\{S_j\}$ because that is the time at which all $m$ coupon types will have been collected. By Theorem 7, $S_j$ has probability density function $f_{S_j}(t) = p_j e^{-p_j t}$, and each $S_j$ is independent. Let us compute the distribution function of $S$:

$$\begin{aligned}
\mathbb{P}(S \le t) &= \mathbb{P}(\max\{S_j\} \le t) \\
&= \mathbb{P}(S_j \le t, \forall j \le m) \\
&= \prod_{j=1}^{m} \mathbb{P}(S_j \le t) \qquad \text{(by independence)} \\
&= \prod_{j=1}^{m} (1 - \mathbb{P}(S_j > t)) \\
&= \prod_{j=1}^{m} (1 - e^{-p_j t}).
\end{aligned}$$

It is a standard fact that for a nonnegative random variable Y, $\mathbb{E}[Y] = \int_0^\infty (1 - F_Y(y))dy$; thus, $\mathbb{E}[S] = \int_0^\infty (1 - F_S(t))dt = \int_0^\infty [1 - \prod_{j=1}^m (1 - e^{-p_j t})]dt$. At this point, we have calculated the expected *time* until all $m$ types of coupons have been collected, not the expected *number* of coupons we would need to collect. Let $I_i$ be the $i$th inter arrival time of coupons. Then clearly $\sum_{i=1}^M I_i = S$, so $\mathbb{E}[\sum_{i=1}^M I_i] = \mathbb{E}[S]$, but since the $I_i$'s are exponential with parameter $\lambda = 1$ and independent and also independent of $M$, we have that $\mathbb{E}[\sum_{i=1}^M I_i] = \mathbb{E}[M]\mathbb{E}[I_i] = \mathbb{E}[M] \cdot 1 = \mathbb{E}[M]$. Thus, the value we computed for $\mathbb{E}[S]$ is also the value of $\mathbb{E}[M]$.

4.3. **Customers Purchasing a Commodity.** Suppose that customers arrive at an amusement park according to a Poisson process of intensity $\lambda$; each customer must pay \$1 to enter the park. If the value of the price of admission is discounted back to time $t = 0$ according to an exponential distribution with rate $\beta$, we can determine the total expected value of money collected; call this value $M$. If $S_1, S_2, \dots$ are the arrival times, and $N(t)$ is the number of people admitted up to time $t$, the expected value of money collected up to time $t$ is given by $M = \mathbb{E}\left[\sum_{k=1}^{N(t)} e^{-\beta S_k}\right]$. To determine this expected value, we use the law of total expectation: We have

$$M = \sum_{n=1}^{\infty} \mathbb{E}\left[\sum_{k=1}^{n} e^{-\beta S_k} | N(t) = n\right] \mathbb{P}(N(t) = n).$$

Recall that if $U_1, U_2, \dots$ are independent and uniformly distributed on the interval $(0, t]$, then the conditional distribution given that $N(t) = n$ of $S_1, S_2, \dots, S_n$ is the same as that of the order statistics of $U_1, U_2, \dots, U_n$, so by the fact that $\sum_{k=1}^{n} e^{-\beta U_k} = \sum_{k=1}^{n} e^{-\beta U_{(k)}}$, we have

$$\begin{aligned}
\mathbb{E}\left[\sum_{k=1}^{N(t)} e^{-\beta S_k} | N(t) = n\right] &= \mathbb{E}\left[\sum_{k=1}^{n} e^{-\beta U_{(k)}}\right] \\
&= \mathbb{E}\left[\sum_{k=1}^{n} e^{-\beta U_k}\right] \\
&= n\mathbb{E}\left[e^{-\beta U_1}\right] \\
&= \frac{n}{t}\int_0^t e^{-\beta u} du \\
&= \frac{n}{\beta t}[1 - e^{-\beta t}].
\end{aligned}$$

Thus, $M = \sum_{n=1}^{\infty} \frac{n}{\beta t}[1 - e^{-\beta t}]\mathbb{P}(N(t) = n) = \frac{1 - e^{-\beta t}}{\beta t} \sum_{n=0}^{\infty} n\mathbb{P}(N(t) = n) = \frac{1 - e^{-\beta t}}{\beta t}\mathbb{E}[N(t)] = \frac{\lambda(1 - e^{-\beta t})}{\beta}$.

4.4. **Shot Noise.** The following example was adapted from [4, p. 326]: Suppose that electrons arrive at the start of a wire according to a Poisson process of rate $\lambda$; $t$ seconds after the arrival of an electron, the current created by that electron is $I(t) = e^{-\beta t}$ for some $\beta > 0$. Let $t > 0$, $N(t) = n$, and suppose arrivals of electrons occur at $S_1, S_2, \ldots, S_n$, then the current in the wire is $X(t) = \displaystyle\sum_{i=1}^{N(t)} I(t - S_i)$. We can now determine the expected value of the current in the wire at time $t$. Let $s > 0$ and let $U_1, U_2, \ldots, U_n$ be uniformly distributed on $(0, t]$, then consider the Laplace transform of $X(t)$:

$$
\begin{aligned}
\mathbb{E}\left[e^{-sX(t)}\right] &= \mathbb{E}\left[e^{-s\sum_{i=1}^{N(t)} I(t-S_i)}\right] \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{-s\sum_{i=1}^{N(t)} I(t-S_i)} | N(t) = n\right] \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{-s\sum_{i=1}^{n} I(t-U_{(i)})}\right] \mathbb{P}(N(t) = n) \qquad \text{(by Proposition 2)} \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{-s\sum_{i=1}^{n} I(t-U_i)}\right] \mathbb{P}(N(t) = n) \qquad \text{(by symmetry of the sum)} \\
&= \sum_{n=0}^{\infty} \mathbb{E}\left[e^{-snI(t-U_1)}\right] \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \left[\mathbb{E}[e^{-sI(t-U_1)}]\right]^n \mathbb{P}(N(t) = n).
\end{aligned}
$$

Recall that $U_1$ is uniformly distributed on $(0,t]$, so $U_1$ and $t-U_1$ have the same distribution. Thus,

$$\begin{aligned}
\sum_{n=0}^{\infty}\left[\mathbb{E}[e^{-sI(t-U_1)}]\right]^n \mathbb{P}(N(t)=n) &= \sum_{n=0}^{\infty}\left[\mathbb{E}[e^{-sI(U_1)}]\right]^n \mathbb{P}(N(t)=n) \\
&= \sum_{n=0}^{\infty}\left[\int_0^t \frac{1}{t}\exp\{-sI(x)\}\,dx\right]^n \mathbb{P}(N(t)=n) \\
&= \sum_{n=0}^{\infty}\left[\int_0^t \frac{1}{t}\exp\{-se^{-\beta x}\}\,dx\right]^n \frac{e^{-\lambda t}(\lambda t)^n}{n!} \\
&= e^{-\lambda t}\sum_{n=0}^{\infty}\left[\lambda\int_0^t \exp\{-se^{-\beta x}\}\,dx\right]^n \frac{1}{n!} \\
&= \exp\left\{\lambda\left[\int_0^t \exp\left(-se^{-\beta x}\right)dx - t\right]\right\} \\
&= \exp\left\{\lambda\int_0^t (\exp\left\{-se^{-\beta x}\right\}-1)dx\right\}.
\end{aligned}$$

Computing the $n$th derivative of the Laplace transform with respect to $s$, setting $s=0$, and multiplying by a factor of $(-1)^n$ yields the $n$th moment of the random variable. Thus, $\mathbb{E}(X(t)) = \lambda\int_0^t e^{-\beta U_1}dU_1 = \frac{\lambda}{\beta}(1-e^{-\beta t})$, so as $t\to\infty$, we have that $\mathbb{E}(X(t)) \nearrow \frac{\lambda}{\beta}$. Other, more general, functions for $I(t)$ can also be considered.

## 5. The Compound Poisson Process

Let $N(t)$ be a Poisson process of intensity $\lambda > 0$, and suppose that each occurrence of the Poisson process has an associated random variable $Y_k$. Further, assume that $Y_1, Y_2, \ldots$ are independent, identically distributed random variables with a common distribution function

$G(y) = \mathbb{P}(Y_k \leq y)$. A **compound Poisson process** is the process

$$Z(t) = \sum_{k=1}^{N(t)} Y_k,$$

where $t > 0$.

Let $G^{(n)}$ denote the distribution function of $Y_1+Y_2+\cdots+Y_n$, for $n = 1, 2, \ldots$, with $G^{(0)}(y) = 1$ for $y \geq 0$, then $G^{(n)}(y) = \mathbb{P}(\sum_{i=1}^{n} Y_i \leq y) = \int_{-\infty}^{\infty} G^{(n-1)}(y-z)\, dG(z)$. With this, we can easily compute the distribution function for $Z(t)$:

$$\begin{aligned}
\mathbb{P}(Z(t) \leq z) &= \mathbb{P}\left(\sum_{i=1}^{N(t)} Y_i \leq z\right) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{N(t)} Y_i \leq z | N(t) = n\right) \mathbb{P}(N(t) = n) \\
&= \sum_{n=0}^{\infty} \mathbb{P}\left(\sum_{i=1}^{n} Y_i \leq z\right) \frac{(\lambda t)^n e^{-\lambda t}}{n!} \\
&= \sum_{n=0}^{\infty} G^{(n)}(z) \frac{(\lambda t)^n e^{-\lambda t}}{n!}.
\end{aligned}$$

## 6. Conclusion

We have seen many different properties of and results related to the Poisson process here. Several of them were straightforward calculations, but many of the proofs and examples presented required insight beyond just computation. An important tool that was used many times was the Law of Total Probability and conditioning of random variables, which often offer a convenient method for calculating probabilities that would otherwise be difficult to compute.

Clearly, the Poisson process has many far-reaching applications, and this thesis only scratches the surface of ways that the Poisson process can be used. In addition to the compound Poisson process, there are other related processes including the spatial Poisson process, whose indexing set $T$ is not $(0, \infty) \subset \mathbb{R}$, but rather $\mathbb{R}^n$, and the marked Poisson process, which considers ordered pairs of random variables; one can also simulate Poisson processes. I hope that this thesis provides a strong enough foundational knowledge that the interested reader could find these additional topics accessible.

## 7. References


(1) Apostol, Tom M. Mathematical Analysis. Reading, MA: Addison-Wesley, 1974. Print.

(2) Ghahramani, Saeed. Fundamentals of Probability with Stochastic Processes. 3rd ed. Upper Saddle River, NJ: Pearson/Prentice Hall, 2005. Print.

(3) Pinsky, Mark A., and Samuel Karlin. An Introduction to Stochastic Modeling. 3rd ed. Amsterdam: Elsevier/Academic, 1998. Print.

(4) Resnick, Sidney I. Adventures in Stochastic Processes. Boston: Birkhuser, 1992. Print.

(5) Ross, Sheldon M. Introduction to Probability Models. 7th ed. San Diego: Academic, 2000. Print.

(6) Rudin, Walter. Real and Complex Analysis. 3rd ed. N.p.: McGraw-Hill, 1986. Print.

(7) Siegrist, Kyle. "13. The Poisson Process." The Poisson Process. University of Alabama at Huntsville, 2013. Web. 27 Mar. 2013.