DMS625: Introduction to stochastic processes and their applications

Sourav Majumdar

Poisson Processes

October 6, 2024

1 Exponential Distribution

Before we begin discussing Poisson processes we will collect some facts about the exponential distribution. A random variable X is said to be **exponentially distributed** with **rate** λ if,

$$\mathbb{P}(X \leq x) = 1 - e^{-\lambda x}, 0 \leq x < \infty$$

We will denote this as $X \sim \text{Exp}(\lambda)$. The density of X is given by,

$$f_X(x) = \lambda e^{-\lambda x}, 0 \leq x < \infty$$

.

Exercises

1. Show that $\mathbb{E}[X] = \frac{1}{\lambda}$
2. Show that $\text{Var}[X] = \frac{1}{\lambda^2}$

The exponential distribution is the only continuous probability distribution to have the **memoryless property**, formally this is given by,

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

Verify that the Exponential distribution satisfies the above property. Suppose the waiting time for a bus, denoted by the random variable X, follows the exponential distribution. Then from the memoryless property $\mathbb{P}(X > 10 + 5 | X > 10) = \mathbb{P}(X > 5)$. This means that suppose you have waited for the bus for 10 minutes, the probability that you have to wait for 5 more minutes is the same as not having waited for the bus at all.

Proposition 1.1. *Let $S \sim Exp(\lambda)$ and $T \sim Exp(\mu)$. S and T are independent. Then,*

1.

$$\mathbb{P}(\min(S, T) > t) = e^{-(\lambda+\mu)t}$$

2.

$$\mathbb{P}(S < T) = \frac{\lambda}{\lambda + \mu}$$

Proof. 1.

$$\begin{aligned}\mathbb{P}(\min(S,T) > t) &= \mathbb{P}(S > t, T > t) \\ &= \mathbb{P}(S > t)\mathbb{P}(T > t) \text{ (From Independence)} \\ &= e^{-\lambda t} e^{-\mu t} = e^{-(\lambda+\mu)t}\end{aligned}$$

2.

$$\begin{aligned}\mathbb{P}(S < T) &= \int_0^\infty \int_s^\infty \lambda e^{-\lambda s} \mu e^{-\mu t} dt ds \\ &= \int_0^\infty \lambda e^{-\lambda s} e^{-\mu s} ds \\ &= \frac{\lambda}{\lambda + \mu}\end{aligned}$$

□

Example 1.1. *Anne and Betty enter a beauty parlor simultaneously. Anne to get a manicure and Betty to get a haircut. Suppose the time for a manicure (haircut) is exponentially distributed with mean 20 (30) minutes.*

1. *What is the probability Anne gets done first?*

   $A \sim Exp(\frac{1}{20\ min}) = Exp(3/hour)$, *i.e.*, $\lambda_A = 3$. *Similarly,* $\lambda_B = 2$.

   $$\mathbb{P}(A < B) = \frac{3}{3+2} = \frac{3}{5}$$

2. *What is the expected amount of time until Anne and Betty are both done?*

   $\mathbb{P}(A < x, B < x) = \mathbb{P}(A < x)\mathbb{P}(B < x)$ *(From Independence)* $= (1 - e^{-3x})(1 - e^{-2x}) = 1 - e^{-3x} - e^{-2x} + e^{-5x}$. *To get the joint density of* A, B *finishing by time* x, *we differentiate* $\mathbb{P}(A < x, B < x)$ *wrt* x.

   $$f_{A,B}(x) = \frac{d}{dx}\left(1 - e^{-3x} - e^{-2x} + e^{-5x}\right) = 3e^{-3x} + 2e^{-2x} - 5e^{-5x}$$

   $$\begin{aligned}\mathbb{E}[A,B \textit{ are both done}] &= \int_0^\infty x f_{A,B}(x) dx \\ &= \int_0^\infty (3xe^{-3x} + 2xe^{-2x} - 5xe^{-5x}) dx \\ &= \int_0^\infty 3xe^{-3x} dx + \int_0^\infty 2xe^{-2x} dx - \int_0^\infty 5xe^{-5x} dx\end{aligned}$$

   Notice that each term represents the expectation of an exponential random variable with rate $2, 3, 5$ *respectively. Therefore,*

   $$\mathbb{E}[A,B \textit{ are both done}] = \frac{1}{3} + \frac{1}{2} - \frac{1}{5} = \frac{19}{30} \textit{ hours}$$

Proposition 1.2. *Let* $T_i \sim Exp(\lambda_i)$ *and* $V = \min(T_1, \ldots, T_n)$. *Then,*

1. $\mathbb{P}(V > t) = \exp(-\sum_i \lambda_i t)$

2. *Let I denote the index of T's that is the smallest in V. $\mathbb{P}(I = i) = \frac{\lambda_i}{\sum_i \lambda_i}$, is the probability that T_i is the minimum in V.*

Proof. 1. Same idea as Proposition 1.1 part 1. Left as an exercise.

2. Let $T = \min(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)$. Then from part 1), $T \sim \text{Exp}(T_1 + \ldots + T_{i-1} + T_{i+1} + \ldots + T_n)$. Now let $S = T_i$. The probability that we are interested in is $\mathbb{P}(T_i = \min(T_1, \ldots, T_n))$, this can be written as,

$$\begin{aligned}
\mathbb{P}(T_i = \min(T_1, \ldots, T_n)) &= \mathbb{P}(T_i < \min(T_1, \ldots, T_{i-1}, T_{i+1}, \ldots, T_n)) \\
&= \mathbb{P}(S < T) \\
&= \frac{\lambda_i}{\sum_i \lambda_i} \text{ (From Proposition 1.1 part 2)}
\end{aligned}$$

□

Example 1.2. *Ram, Shyam and Ghanshyam arrive at a government office. The amount of time they will stay is exponentially distributed with means of $1, 1/2$ and $1/3$ hours respectively. For each find the probability that they are the last ones left in the office.*

Let X_0 be the random variable that denotes the set of people in the office at the start, here X_0 can only take the value RSG, i.e. $\mathbb{P}(X_0 = RSG) = 1$. Let X_1 denote the set of people remaining after the first one has left. X_1 can take values in $\{RS, RG, SG\}$. Similarly X_2 denotes the set of people in the office remaining after two of them have left, X_2 can take values in $\{R, S, G\}$. Let T_R, T_S, T_G denote the time they take to leave.

Then,

$$\begin{aligned}
\mathbb{P}(X_2 = R) &= \mathbb{P}(X_2 = R \cap \{\{X_1 = RS, X_0 = RSG\} \cup \{X_1 = RG, X_0 = RSG\} \cup \{X_1 = SG, X_0 = RSG\}\}) \\
&= \mathbb{P}(X_2 = R \cap \{X_1 = RS, X_0 = RSG\}) + \mathbb{P}(X_2 = R \cap \{X_1 = RG, X_0 = RSG\}) \\
&+ \mathbb{P}(X_2 = R \cap \{X_1 = SG, X_0 = RSG\}) \\
&\textit{(Follows from the axiom that prob. of union of disjoint events is the sum of the prob. of the events)} \\
&= \mathbb{P}(X_2 = R | X_1 = RS, X_0 = RSG)\mathbb{P}(X_1 = RS | X_0 = RSG)\mathbb{P}(X_0 = RSG) \\
&+ \mathbb{P}(X_2 = R | X_1 = RG, X_0 = RSG)\mathbb{P}(X_1 = RG | X_0 = RSG)\mathbb{P}(X_0 = RSG) \\
&+ \mathbb{P}(X_2 = R | X_1 = SG, X_0 = RSG)\mathbb{P}(X_1 = SG | X_0 = RSG)\mathbb{P}(X_0 = RSG) \\
&= \mathbb{P}(X_2 = R | X_1 = RS)\mathbb{P}(X_1 = RS | X_0 = RSG) + \mathbb{P}(X_2 = R | X_1 = RG)\mathbb{P}(X_1 = RG | X_0 = RSG) \\
&+ \mathbb{P}(X_2 = R | X_1 = SG)\mathbb{P}(X_1 = SG | X_0 = RSG) \\
&\textit{(Note that } \mathbb{P}(X_2 = R | X_1 = RS, X_0 = RSG) = \mathbb{P}(X_2 = R | X_1 = RS), \textit{ since } \mathbb{P}(X_0 = RSG) = 1.\textit{)} \\
&= \mathbb{P}(T_S < T_R)\mathbb{P}(T_G < \min(T_R, T_S)) + \mathbb{P}(T_G < T_R)\mathbb{P}(T_S < \min(T_R, T_G)) \\
&(\mathbb{P}(X_2 = R | X_1 = SG) = 0 \textit{ and } \mathbb{P}(X_1 = RS | X_0 = RSG) = \mathbb{P}(T_G < \min(T_R, R_S), \\
&\textit{similar interpretations follow for other terms)} \\
&= \frac{\lambda_S}{\lambda_R + \lambda_S} \frac{\lambda_G}{\lambda_R + \lambda_S + \lambda_G} + \frac{\lambda_G}{\lambda_R + \lambda_G} \frac{\lambda_S}{\lambda_R + \lambda_S + \lambda_G} \\
&\textit{(From Proposition 1.2 part 2)} \\
&= \left(\frac{2}{2+1}\right)\left(\frac{3}{1+2+3}\right) + \left(\frac{2}{1+2+3}\right)\left(\frac{3}{1+3}\right) = \frac{7}{12}
\end{aligned}$$

Similarly calculate for others.

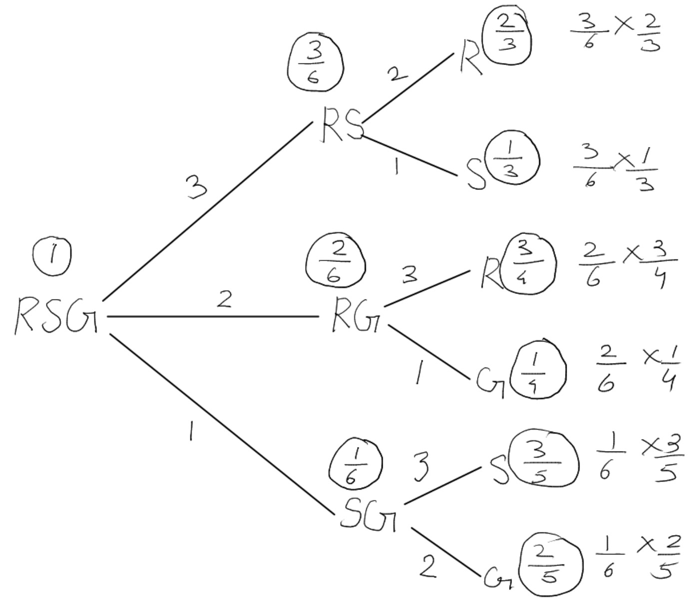

Figure 1: The figure represents the pathwise probabilities for each individual to be the last one remaining. The formal argument for the calculation is given above. The numbers on the line denotes the rate of the individual that left. The numbers in circle denote the probabilities.

2 Poisson Distribution

We say that X follows the **Poisson distribution** with mean λ if,

$$\mathbb{P}(X = n) = e^{-\lambda}\frac{\lambda^n}{n!}, n = 0, 1, 2, \ldots$$

We denote it as $X \sim \text{Poi}(\lambda)$.

Exercises

1. Show that $\mathbb{E}[X] = \lambda$
2. Show that $\text{Var}[X] = \lambda$

Proposition 2.1. *Let $X_i \sim Poi(\lambda_i)$ and are independent. Then,*

$$X_1 + \ldots + X_k \sim Poi(\lambda_1 + \ldots + \lambda_k)$$

Proof.

$$\begin{aligned}
\mathbb{P}(X_1 + X_2 = n) &= \sum_{m=0}^{n} \mathbb{P}(X_1 = m, X_2 = n-m) \\
&= \sum_{m=0}^{n} \mathbb{P}(X_1 = m)\mathbb{P}(X_2 = n-m) \\
&= \sum_{m=0}^{n} \frac{e^{-\lambda_1^m}}{m!} \frac{e^{-\lambda_2^{n-m}}}{(n-m)!} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!} \sum_{m=0}^{n} \binom{n}{m} \left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^m \left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-m} \\
&= \frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!} \cdot 1
\end{aligned}$$

Similarly, argue for the higher cases using induction. □

To interpret the Poisson distribution and motivate subsequently the definition of Poisson process, we consider the following result.

Theorem 2.1. *If n is large then, Bin$(n, \lambda/n)$ is approximately Poi(λ).*

Proof. Consider $\lim_{n\to\infty} \mathbb{P}(X = m)$, where $X \sim \text{Bin}(n, \lambda/n)$ and let $Y \sim \text{Poi}(\lambda)$.

$$\begin{aligned}
\lim_{n\to\infty} \mathbb{P}(X = m) &= \lim_{n\to\infty} \binom{n}{m} \left(\frac{\lambda}{n}\right)^m \left(1 - \frac{\lambda}{n}\right)^{n-m} \\
&= \lim_{n\to\infty} \frac{\lambda^m}{m!} \frac{n(n-1)\dots(n-m+1)}{n^m} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-m} \\
&= \frac{\lambda^m}{m!} \cdot 1 \cdot e^{-\lambda} \cdot 1 \\
&= \mathbb{P}(Y = m)
\end{aligned}$$

□

Note that if $X \sim \text{Bin}(n, \lambda/n)$ then $\mathbb{E}[X] = \lambda$. Even though $n \to \infty$, the expected number of successes remains λ. The theorem states that if there are infinitely many trials and the expected number of successes stays λ, then the Poisson distribution approximates the binomial distribution. Since the binomial distribution gives the probability of m successes in n trials, as the number of trials becomes large the probability of m successes in these large number of trials is given by the Poisson distribution.

3 Poisson Process

Poisson process is a particular type of a **counting process**. A counting process, as the name implies, counts the occurrence of certain events. If $N(t)$ is a counting process, that counts the number of calls arriving at a call centre. Then $N(t)$ is the number of calls that have arrived in the call centre until time t has passed.

We say that $\{N(s), s \geq 0\}$ is a **Poisson process**, if,

1. $N(0) = 0$
2. $N(t+s) - N(s) \sim \text{Poi}(\lambda t)$

3. $N(t)$ has **independent increments**, i.e., $t_0 < t_1 < \ldots < t_n$, then, $N(t_1) - N(t_0), \ldots, N(t_n) - N(t_{n-1})$ are independent.

Let's examine each of the properties for the Poisson process. Let's say the Poisson process is being used to model calls arriving at a call centre. The first condition says that at $t = 0$, the count of the number of calls arrived is 0. The second condition says that in a time interval $(s, s+t]$, the number of calls that arrive is distributed as $\text{Poi}(\lambda t)$, the length of the time interval is t. Note the second condition implies that the number of arrivals depends only on the length of the interval, i.e., time intervals $(0, 5]$ and $(50, 55]$ will have the same distribution of the number of arrivals. The third condition says that the number of arrivals in disjoint time intervals are independent, i.e., the number of calls that arrive in $(t_0, t_1], (t_1, t_2], \ldots, (t_{n-1}, t_n]$ are mutually independent.

Exercises

1. Verify that property 2 and property 3 for the Poisson process are consistent, i.e., property 3 follows from property 2.

Poisson processes find applications in many places.

- In the early 1900s to model the number of calls that arrive at a telephone exchange, Erlang proposed to model how the calls arrive at the exchange using Poisson processes. This was necessary so as to be able to have sufficient telephone operators to be able to manage demand. This led to the birth of a more general subject called Queuing theory and is a major area in Operations Research.
- Poisson processes are used to model/count the number of *shocks* received by a system, say the failures of a component, natural disasters, shocks to economy/market, etc.
- Queuing theory more broadly finds application in managing traffic, planning layouts to handle queues in a warehouse. More contemporary applications would include modelling limit order books in financial markets.

Exercises

1. Show that $\mathbb{P}(N(t) = n) = e^{-\lambda t} \frac{(\lambda t)^n}{n!}$
2. $\mathbb{E}[N(t)] = \lambda t$
3. $\text{Var}[N(t)] = \lambda t$

Interarrival time

Let X_n denote the **interarrival time** between the n-th and $(n-1)$-th event, i.e. the time gap between the n-th and the $(n-1)$-th event.

Then X_1 denotes the time of the first event. If the first event happens after time t, then that implies that there were no events between $(0, t]$. Therefore,

$$\begin{aligned}
\mathbb{P}(X_1 > t) &= \mathbb{P}(N(t) = 0) \\
&= \mathbb{P}(N(t) - N(0) = 0) \\
&\text{(From the first property } N(0) = 0 \text{ and from the second property of Poisson process,} \\
&N(t) - N(0) \sim \text{Poi}(\lambda t)) \\
&= e^{-\lambda t} \frac{(\lambda t)^0}{0!} \\
&= e^{-\lambda t}
\end{aligned}$$

Therefore, $X_1 \sim \text{Exp}(\lambda)$. Similarly, observe that X_2 is the time between the first and the second event. Say the first event happened at some time s, then $\mathbb{P}(X_2 > t)$ denotes that that no event happened between $(s, s+t]$. Therefore,

$$\mathbb{P}(X_2 > t) = \mathbb{P}(N(s+t) - N(s) = 0) = e^{-\lambda t}\frac{(\lambda t)^0}{0!} = e^{-\lambda t}$$

Therefore, $X_2 \sim \text{Exp}(\lambda)$. Now consider,

$$\begin{aligned}
\mathbb{P}(X_2 > t | X_1 = s) &= \mathbb{P}(0 \text{ events in } (s, s+t] | X_1 = s) \\
&= \mathbb{P}(0 \text{ events in } (s, s+t] | 1 \text{ event in } (0, s]) \\
&= \mathbb{P}(0 \text{ events in } (s, s+t]) \\
&\text{(By independent increments property)} \\
&= \mathbb{P}(N(t+s) - N(s) = 0) \\
&= e^{-\lambda t}
\end{aligned}$$

Hence, X_2 is independent of X_1. Similarly, we can conclude the following general result,

Theorem 3.1. $X_1, \ldots, X_n$ *are independent and identically distributed as Exp*(λ).

Remark 1. *Note that* X_n*, the interarrival time, is* ***not*** *the time till the* n*-th event, but rather the time between the* n*-th and* $(n-1)$*-th event. To compute the time till the* n*-th event we have to evaluate the waiting time which is considered next.*

Waiting time

The **waiting time**, S_n, for the n-th event to arrive is,

$$S_n = \sum_{i=1}^{n} X_i$$

$$\begin{aligned}
\mathbb{P}(S_n \leq t) &= \mathbb{P}(N(t) \geq n) \\
&\quad (S_n \leq t \text{ means that } n \text{ events have arrived before time } t\text{, this can be written as } N(t) \geq n) \\
&= \sum_{j=n}^{\infty} e^{-\lambda t}\frac{(\lambda t)^j}{j!}
\end{aligned}$$

We next want to obtain the probability density function, $f_{S_n}(t)$, of S_n,

$$\begin{aligned}
f_{S_n}(t) &= \frac{d}{dt}\mathbb{P}(S_n \leq t) \\
&= \frac{\lambda^n e^{-\lambda t} t^{n-1}}{(n-1)!}
\end{aligned}$$

This is a very well known density function of the Gamma distribution. Since n is restricted to be an integer here, this special case is called the Erlang distribution.

Exercises

1. Verify that $\mathbb{E}[S_n] = \frac{n}{\lambda}$.
2. Verify that $\text{Var}[S_n] = \frac{n}{\lambda^2}$.

Example 3.1. *Show that* $\mathbb{P}(N(s) = k | N(t) = n) = \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}; s < t.$

$$
\begin{aligned}
\mathbb{P}(N(s) = k | N(t) = n) &= \frac{\mathbb{P}(N(s) = k, N(t) = n)}{\mathbb{P}(N(t) = n)} \\
&= \frac{\mathbb{P}(k \textit{ arrivals in } (0, s],\ n \textit{ arrivals in } (0, t])}{\mathbb{P}(N(t) = n)} \\
&= \frac{\mathbb{P}(k \textit{ arrivals in } (0, s],\ n\text{-}k \textit{ arrivals in } (s, t])}{\mathbb{P}(N(t = n))} \\
&= \frac{\mathbb{P}(N(s) - N(0) = k)\mathbb{P}(N(s + (t - s)) - N(s) = n - k)}{\mathbb{P}(N(t) = n)} \\
&\textit{(By independent increments property)} \\
&= \frac{\frac{e^{-\lambda s}(\lambda s)^k}{k!} \frac{e^{-\lambda(t-s)}(\lambda(t-s))^{n-k}}{(n-k)!}}{\frac{e^{-\lambda t}(\lambda t)^n}{n!}} \\
&= \binom{n}{k} \left(\frac{s}{t}\right)^k \left(1 - \frac{s}{t}\right)^{n-k}
\end{aligned}
$$

Example 3.2. *Calculate* $\mathbb{E}[N(t)N(t + s)]$.

If A, B are independent events then $\mathbb{E}[AB] = \mathbb{E}[A]\mathbb{E}[B]$. Now note here that $N(t), N(t + s)$ are not independent events. But we know that $N(t) - N(0), N(t + s) - N(t)$ are independent events by the independent increments property. This is the idea behind the subsequent calculations.

$$
\begin{aligned}
\mathbb{E}[N(t)N(t + s)] &= \mathbb{E}[N(t)\{N(t + s) - N(t) + N(t)\}] \\
&= \mathbb{E}[N(t)\{N(t + s) - N(t)\}] + \mathbb{E}[N^2(t)] \\
&= \mathbb{E}[N(t)]\mathbb{E}[N(t + s) - N(s)] + \mathbb{E}[N^2(t)] \\
&\textit{(Follows from independence of } N(t) \textit{ and } N(t + s) - N(s)\textit{)} \\
&= (\lambda t)(\lambda s) + \textit{Var}[N(t)] + (\mathbb{E}[N(t)])^2 \\
&= (\lambda t)(\lambda s) + \lambda t + (\lambda t)^2
\end{aligned}
$$

Superposition

If you have a stream of independent incoming Poisson processes, it turns out that they add up to also form a Poisson process.

Theorem 3.2. *Suppose $\{N_1(t), t \geq 0\}$ and $\{N_2(t), t \geq 0\}$ are independent Poisson processes with rates λ_1 and λ_2. Let $N(t) = N_1(t) + N_2(t)$. Then $\{N(t), t \geq 0\}$ is a Poisson process with rate $\lambda_1 + \lambda_2$.*

Proof. The first two properties are left as an exercise. To verify the third property consider the intervals $0 < t_0 < t_1$. Let $N_1(t_0) = k, N_2(t_0) = l, N_1(t_1) - N_1(t_0) = m, N_2(t_1) - N_2(t_0) = n$, then,

$$
\begin{aligned}
&\mathbb{P}(N(t_0) = k + l, N(t_1) - N(t_0) = m + n) \\
=&\mathbb{P}((N_1(t_0) + N_2(t_0)) = k + l, (N_1(t_1) - N_1(t_0)) + (N_2(t_1) - N_2(t_0)) = m + n) \\
=&\mathbb{P}(k \text{ arrivals in } N_1 \text{ in } (0, t_0], l \text{ arrivals in } N_2 \text{ in } (0, t_0], \\
&m \text{ arrivals in } N_1 \text{ in } (t_0, t_1], n \text{ arrivals in } N_2 \text{ in } (t_0, t_1]) \\
&\text{(Now note that events highlighted in red and orange are independent,} \\
&\text{since } N_1(t) \text{ and } N_2(t) \text{ are independent Poisson processes)} \\
=&\mathbb{P}(k \text{ arrivals in } N_1 \text{ in } (0, t_0], l \text{ arrivals in } N_2 \text{ in } (0, t_0])\mathbb{P}(m \text{ arrivals in } N_1 \text{ in } (t_0, t_1], n \text{ arrivals in } N_2 \text{ in } (t_0, t_1]) \\
=&\mathbb{P}(N(t_0) = k + l)\mathbb{P}(N(t_1) - N(t_0) = m + n)
\end{aligned}
$$

□

A more general result also holds,

Corollary 3.1. *If $\{N_i(t), t \geq 0\}$ are independent Poisson processes with rate λ_i. Then $N(t) = \sum_{i=1}^{n} N_i(t)$ is a Poisson process with rate $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.*

The implication is that if we add n independent Poisson processes, the sum will also be a Poisson process.

Thinning

The main result of this section is, if you have a Poisson process stream and you split it into several other streams, the splits also are independent Poisson process. Formally consider the setup that $N(t)$ is a Poisson process with rate λ. Now you choose to split $N(t)$ into some j streams. Suppose at time t, $N(t) = n$, this means that n items will each be sent into one of the j streams. Let $Y_i, 1 \leq i \leq n$, denote the random variable measuring the stream that item i has been sent to and $\mathbb{P}(Y_i = j)$ denotes the probability of item i to be sent to stream j. Also $N(t) = \sum_j N_j(t)$.

To explain the notation, suppose people arrive at a mall with a Poisson process, $N(t)$ and rate λ. Now when people leave the mall, they can choose to either opt to travel in a *car* or *bike* ($j = \{bike\}$ or $\{car\}$). They can choose to travel by a *car* with probability 0.4 and by a *bike* with probability 0.6. Therefore $\mathbb{P}(Y_i = car) = 0.4$ and $\mathbb{P}(Y_i = bike) = 0.6$, $1 \leq i \leq N(t)$.

Theorem 3.3. *$N_j(t)$ are independent Poisson processes with rate $\lambda\mathbb{P}(Y_i = j)$.*

Proof. First consider the case that $N(t)$ is split into $m = 2$ streams $N_1(t)$ and $N_2(t)$. The probability of going into each is p and $(1 - p)$. It easily follows that $N_1(0) = 0$ and $N_2(0) = 0$.

Consider the increments $X_1 = N_1(t + s) - N_1(s)$ and $X_2 = N_2(t + s) - N_2(s)$. We will next evaluate their joint distribution,

$$
\begin{aligned}
&\mathbb{P}(X_1 = k, X_2 = l) \\
=&\mathbb{P}(X_1 = k, N(t + s) - N(s) = k + l) \\
=&\mathbb{P}(X_1 = k | N(t + s) - N(s) = k + l)\mathbb{P}(N(t + s) - N(s) = k + l) \\
&\text{(Notice that the term in blue represents the probability of } k \text{ successes in } k + l \text{ independent trials)} \\
=&\binom{k + l}{k} p^k (1 - p)^l e^{-\lambda t} \frac{(\lambda t)^{k+l}}{(k + l)!} \\
=&e^{-\lambda p t} \frac{(\lambda p t)^k}{k!} e^{-\lambda(1-p)t} \frac{(\lambda(1 - p)t)^l}{l!}
\end{aligned}
$$

Therefore, $X_1 \sim \text{Poi}(\lambda p t)$ and $X_2 \sim \text{Poi}(\lambda(1 - p)t)$, and they are independent. Hence, $N_1(t)$ is a Poisson process with rate λp and $N_2(t)$ is a Poisson process with rate $\lambda(1 - p)$, and are independent.

The case for $m > 2$ is left as an exercise. □

Exercises

1. Suppose that people immigrate into a territory according to Poisson process with rate $\lambda = 2$.

    1. Find the probability there are 10 arrivals in the following week.
    2. Find the expected number of days until there have been 20 arrivals.

2. Ellen catches fish at times of a Poisson process with rate 2 per hour. 40% fish are salmon and 60% are trout. What is the probability that she will catch exactly one salmon and two trout if she fishes for 2.5 hours.

Example 3.3. *Two copy editors read a 300-page manuscript. The first found 100 typos and the second found 120 typos. There are 80 common typos. Each editor finds a typo with probability p_1 and p_2. Author makes typos with rate λ per page. Estimate λ, p_1, p_2.*

The key idea here is to identify the right splits to be able to use the information available. The splits are,

$$\begin{aligned}
X_1(t) &= \textit{Typos identified by only Editor 1}\\
X_2(t) &= \textit{Typos identified by only Editor 2}\\
X_3(t) &= \textit{Typos identified by both editors}\\
X_4(t) &= \textit{Typos identifies by none}
\end{aligned}$$

Since the typos are generated by a Poisson process, the splits $X_1(t), X_2(t), X_3(t), X_4(t)$ are independent Poisson processes. The split to X_1 happens with probability $p_1(1-p_2)$, to X_2 with $(1-p_1)p_2$, to X_3 with p_1p_2 and to X_4 with $(1-p_1)(1-p_2)$. Next we will assume that the values observed are the expected values of the system and perform estimation based on that,

$$\begin{aligned}
\mathbb{E}[X_1(300)] &= 300\lambda p_1(1-p_2) = 20\\
\mathbb{E}[X_2(300)] &= 300\lambda(1-p_2)p_1 = 40\\
\mathbb{E}[X_3(300)] &= 300\lambda p_1 p_2 = 80
\end{aligned}$$

From these set of equations one can estimate λ, p_1, p_2.

An alternate definition of Poisson process

We will now consider an alternate definition of the Poisson process. The purpose of this definition will be clear shortly. For this we will require the notion of a function being $\mathbf{o(h)}$.

We say a function $f(x)$ is $o(h)$ if,

$$\lim_{h\to 0} \frac{f(h)}{h} = 0$$

Exercises

1. Verify which of the following functions are $o(h)$

    1. $f(x) = x^2$
    2. $f(x) = \sqrt{x}$
    3. $f(x) = 0$
    4. $f(x) = e^{-x}$

2. Suppose $f_1(x)$ and $f_2(x)$ are $o(h)$, then show that $f_1(x) + f_2(x)$ is also $o(h)$. This result can be informally stated as "$o(h) + o(h) = o(h)$".

3. Suppose $f(x)$ is $o(h)$, then show that $cf(x)$, where c is independent of x is also $o(h)$. Informally this result can be stated as "$c \times o(h) = o(h)$".

We will now state the alternate definition of the Poisson process. The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process with rate $\lambda, \lambda > 0$, if

1. $N(0) = 0$
2. The process has independent increments.
3. $\mathbb{P}(N(t+h) - N(t) = 1) = \lambda h + o(h)$
4. $\mathbb{P}(N(t+h) - N(t) \geq 2) = o(h)$

We will call this as **Definition 2**. We shall call the previous one as **Definition 1**.

Note that it follows (verify!) from Definition 2, that,

$$\mathbb{P}(N(t+h) - N(t) = 0) = 1 - \lambda h + o(h)$$

We will now show that Definition 1 implies Definition 2. This means we have to show that property 2 of Definition 1 implies property 3 and 4 of Definition 2, since rest of the properties are identical.

From Definition 1,

$$\begin{aligned}
\mathbb{P}(N(t+h) - N(t) = 1) &= e^{-\lambda h}\lambda h \\
&= \left(1 - \lambda h + \frac{(\lambda h)^2}{2!} + \dots\right)\lambda h \\
&= \lambda h - (\lambda h)^2 + \frac{(\lambda h)^3}{2!} + \dots \\
&= \lambda h + o(h)
\end{aligned}$$

and,

$$\begin{aligned}
\mathbb{P}(N(t+h) - N(t) \geq 2) &= \sum_{j=2}^{\infty} e^{-\lambda h}\frac{(\lambda h)^j}{j!} \\
&= \sum_{j=2}^{\infty}\left(1 - \lambda h + \frac{(\lambda h)^2}{2!} + \dots\right)\frac{(\lambda h)^j}{j!} \\
&= \sum_{j=2}^{\infty}\left(\frac{(\lambda h)^j}{j!} - \frac{(\lambda h)^{j+1}}{j!} + \dots\right) \\
&= o(h)
\end{aligned}$$

Therefore, Definition 1 implies Definition 2.

We will now show that Definition 2 implies Definition 1. We will define,

$$P_n(t) = \mathbb{P}(N(t+s) - N(s) = n)$$

and hence,

$$P_n(t+h) = \mathbb{P}(N(t+s+h) - N(s) = n)$$

Consider the case for $n = 0$,

$$\begin{aligned}
P_0(t+h) &= \mathbb{P}(N(t+s+h) - N(s) = 0) \\
&= \mathbb{P}(0 \text{ arrivals in } (s, t+s]), 0 \text{ arrivals in } (t+s, t+s+h]) \\
&= \mathbb{P}(0 \text{ arrivals in } (s, t+s]))\mathbb{P}(0 \text{ arrivals in } (t+s, t+s+h]) \\
&= P_0(t)(1 - \lambda h + o(h)) \\
\lim_{h \to 0} \frac{P_0(t+h) - P_0(t)}{h} &= \lim_{h \to 0} \left(-\frac{\lambda h}{h} + \frac{o(h)}{h} \right) P_0(t) \\
&\text{(Rearranging and taking limit)} \\
\frac{d}{dt} P_0(t) &= -\lambda P_0(t)
\end{aligned}$$

Upon solving,

$$P_0(t) = K e^{-\lambda t}$$

The initial condition is $P_0(0) = 1$, follows from $N(0) = 0$ condition for Poisson process and hence $K = 1$. Therefore,

$$P_0(t) = e^{-\lambda t}$$

Now we will consider the case for a general n,

$$\begin{aligned}
P_n(t+h) &= \mathbb{P}(N(t+s+h) - N(s) = n) \\
&= \mathbb{P}(N(t+s) - N(s) = n, N(t+s+h) - N(t+s) = 0) \\
&+ \mathbb{P}(N(t+s) - N(s) = n-1, N(t+s+h) - N(t+s) = 1) \\
&+ \sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \\
&= \mathbb{P}(N(t+s) - N(s) = n)\mathbb{P}(N(t+s+h) - N(t+s) = 0) \\
&+ \mathbb{P}(N(t+s) - N(s) = n)\mathbb{P}(N(t+s+h) - N(t+s) = 1) \\
&+ \sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \\
&= P_n(t)(1 - \lambda h + o(h)) + P_{n-1}(t)(\lambda h + o(h)) + o(h) \\
\lim_{h \to 0} \frac{P_n(t+h) - P_n(t)}{h} &= \lim_{h \to 0} P_n(t)(-\lambda + \frac{o(h)}{h}) + P_{n-1}(t)(\lambda + \frac{o(h)}{h}) + \frac{o(h)}{h} \\
&\text{(Rearranging and taking limit)} \\
\frac{d}{dt} P_n(t) &= -\lambda P_n(t) + \lambda P_{n-1}(t)
\end{aligned}$$

Therefore, from the above ODE and induction, one can show that,

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

Hence, Definition 2 implies Definition 1.

Remark 2. *To understand how the blue terms are equal, note that* $\mathbb{P}(A \cap B) \leq \mathbb{P}(B)$*. Hence,*

$$\begin{aligned}
&\sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \\
\leq &\sum_{k=2}^{\infty} \mathbb{P}(N(t+s+h) - N(t+s) = k) \\
= &\mathbb{P}(N(t+s+h) - N(t+s) \geq 2) \\
= &o(h) \\
&\textit{(Follows from the fourth property of Definition 2 of Poisson process)}
\end{aligned}$$

Therefore,

$$\sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \leq o(h)$$

Also since probabilities are by definition non-negative therefore, we obtain,

$$0 \leq \sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \leq o(h)$$

$$\lim_{h \to 0} \frac{0}{h} \leq \lim_{h \to 0} \frac{1}{h} \sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) \leq \lim_{h \to 0} \frac{o(h)}{h}$$

(Dividing by h *and taking limit)*

$$\lim_{h \to 0} \frac{1}{h} \sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) = 0$$

(By the sandwich theorem)

And hence,

$$\sum_{k=2}^{\infty} \mathbb{P}(N(t+s) - N(s) = n-k, N(t+s+h) - N(t+s) = k) = o(h)$$

Non-homogeneous Poisson process

Consider traffic on highway. Is the distribution of traffic throughout the day identical? Would you expect the flow of traffic between 12:00 am and 1:00 am, to be the same as the flow of traffic between 12:00 pm and 1:00 pm? You wouldn't since usually in the night there is reduced traffic and during the day there is increased traffic. However if you were to model traffic using Poisson process, from the second property of Definition 1, i.e.,

$$\mathbb{P}(N(t+s) - N(s)) \sim \text{Poi}(\lambda t)$$

any interval of length t would have the same probability distribution of traffic. Since 12:00 am-1:00 am, & 12:00 pm - 1:00 pm, are both 1 hour time intervals, the implication of modelling with Poisson process is that the traffic is identically distributed in each of these time intervals. This is clearly incorrect and undesirable from a modelling point of view. So what can we do in such situations?

We note here that this phenomenon is a consequence of rate of arrivals in the night and day being different. So perhaps we may be able to reflect reality better through our model if we can incorporate *rates that vary in time.* This leads us to the non-homogeneous Poisson process, which is an extension of Poisson process with time-varying rates. To define non-homogeneous Poisson process we will modify Definition 2 of the Poisson process with time varying rates.

A counting process $\{N(t), t \geq 0\}$ is said to be a **non-homogeneous Poisson process** with rate $\lambda(t) > 0$ if,

1. $N(0) = 0$
2. $N(t)$ possesses independent increments.
3. $\mathbb{P}(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$
4. $\mathbb{P}(N(t+h) - N(t) \geq 2) = o(h)$

Note here that the rate, $\lambda(t)$, is a function of time. We have modified Definition 2 of Poisson process with rate $\lambda(t)$ to define non-homogeneous Poisson process.

For $N(t)$ a non-homogeneous Poisson process with rate $\lambda(t)$, define,

$$P_n(t) = \mathbb{P}(N(t+s) - N(s) = n)$$

If we follow the procedure of the previous section we can show that,

$$\begin{aligned}
P_0(t+h) &= P_0(t)(1 - \lambda(t+s)h + o(h)) \\
\frac{d}{dt}P_0(t) &= -\lambda(t+s)P_0(t) \\
&\quad \text{(Rearranging and taking limits as was done in the last section)} \\
\frac{1}{P_0(t)}\frac{d}{dt}P_0(t) &= -\lambda(t+s) \\
\frac{d}{dt}\ln P_0(t) &= -\lambda(t+s) \\
P_0(t) &= e^{-\int_0^t \lambda(u+s)du} \\
P_0(t) &= e^{-\int_s^{t+s} \lambda(z)dz} \\
&\quad \text{(Change of variable with } z = u+s\text{)} \\
P_0(t) &= e^{-\left(\int_0^{t+s} \lambda(z)dz - \int_0^s \lambda(z)dz\right)}
\end{aligned}$$

We define, $m(s) = \int_0^s \lambda(z)dz$, therefore,

$$P_0(t) = e^{-(m(t+s)-m(s))}$$

Using the same approach as in the last section we can show the following,

Theorem 3.4. *For a non-homogeneous Poisson process $N(t)$ with rate $\lambda(t)$,*

$$P_n(t) = \mathbb{P}(N(t+s) - N(s) = n) = e^{-(m(t+s)-m(s))}\frac{(m(t+s) - m(s))^n}{n!}$$

i.e., $N(t+s) - N(s) \sim Poi(m(t+s) - m(s))$

Exercises

1. On a highway between 12:00 am to 6:00 am 30 cars pass per hour, between 6:00 am to 12:00 pm 200 cars pass per hour. Model the arrival of cars on highway as non-homogeneous Poisson process. Find the probability that 120 cars passed between 5:00 am and 7:00 am.

2. On a highway cars arrive at a rate of $60 + 50\sin\left(\frac{\pi t}{12}\right)$, here t is denoted in hours. Find the probability that in a span of 24 hours, 1000 cars passed through the highway.

Though non-homogeneous Poisson processes offer us the flexibility of working with time-varying rates, we lose some nice properties of the Poisson process. One such result is the following,

Proposition 3.1. *For a non-homogeneous Poisson process $N(t)$ with rate $\lambda(t)$, interarrival times need not be exponentially distributed and independent.*

Proof. Left as an exercise. □

Therefore, the waiting time distribution for a non-homogeneous process also need not be Erlang distribution.

Sampling view of non-homogeneous Poisson process

It is possible to view any non-homogeneous Poisson process as a sample from a Poisson process. Consider a Poisson process $N(t)$ with rate λ. Now suppose you sample the arrivals of $N(t)$ with probability $\frac{\lambda(t)}{\lambda}$ independently at time t where $\lambda(t) \leq \lambda$. This means that you count an arrival of $N(t)$ with probability $\frac{\lambda(t)}{\lambda}$. It then follows that $N_c(t)$, the arrivals of $N(t)$ that you have counted, is a non-homogeneous Poisson process with rate $\lambda(t)$.

What does this perspective achieve? First, it gives us a probabilistic connection between Poisson process and non-homogeneous Poisson process. We can understand a non-homogeneous Poisson process in terms of a Poisson process, which are simpler to state. Second, it gives us a way to simulate from non-homogeneous Poisson process, by first simulating a Poisson process and then sampling from it.

We will now show that $N_c(t)$ is indeed a non-homogeneous Poisson process with rate $\lambda(t)$. First, we will need a result on the conditional arrival time T of a Poisson process. We are interested in understanding the distribution of the arrival time T of an event given that there was a single arrival in a time interval $(s, t+s]$.

$$\begin{aligned}
\mathbb{P}(T \leq u | N(t+s) - N(s) = 1) &= \frac{\mathbb{P}(N(t+s) - N(s) = 1, T \leq u)}{\mathbb{P}(N(t+s) - N(s) = 1)} \\
&= \frac{\mathbb{P}(1 \text{ event in } (s,u])\mathbb{P}(0 \text{ event in } (u, t+s])}{\mathbb{P}(N(t+s) - N(s) = 1)} \\
&= \frac{e^{-\lambda(u-s)}\lambda(u-s)e^{-\lambda(t+s-u)}}{e^{-\lambda t}(\lambda t)} \\
&= \frac{u-s}{t}
\end{aligned}$$

Therefore the density of the conditional arrival time is,

$$f_{T|N(t+s)-N(s)=1} = \frac{d}{du}\mathbb{P}(T \leq u | N(t+s) - N(s) = 1) = \frac{1}{t}$$

1. $N_c(0) = 0$

   It follows since $N(0) = 0$, it implies that $N_c(0) = 0$ since there are no arrivals to sample from.

2. Independent increments

   Consider two disjoint time intervals $(s, t]$ and $(t, t+h]$. Since $N(t) - N(s)$ and $N(t+h) - N(t)$ are independent, and we sample independently, therefore, $N_c(t) - N_c(s)$ and $N_c(t+h) - N_c(t)$ are independent.

3. $\mathbb{P}(N_c(t+h)-N_c(t)=1)=\lambda(t)h+o(h)$
   First we note that,

$$
\begin{aligned}
&\mathbb{P}(N_c(t+h)-N_c(t)=1|N(t+h)-N(t)=1)\\
=&\int_t^{t+h}\mathbb{P}(N_c(t+h)-N_c(t)=1|N(t+h)-N(t)=1,T=t^*)f_{T|N(t+h)-N(t)=1}(t^*)dt^*\\
=&\int_s^{t+h}\binom{1}{1}\left(\frac{\lambda(t^*)}{\lambda}\right)\frac{1}{h}dt^*\\
=&\frac{1}{\lambda h}\int_t^{t+h}\lambda(t^*)dt^*
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{P}(N_c(t+h)-N_c(t)=1)=&\,\mathbb{P}(N_c(t+h)-N_c(t)=1|N(t+h)-N(t)=1)\mathbb{P}(N(t+h)-N(t)=1)\\
&+\sum_{k=2}^{\infty}\mathbb{P}(N_c(t+h)-N_c(t)=1|N(t+h)-N(t)=k)\mathbb{P}(N(t+h)-N(t)=k)\\
=&\,\frac{1}{\lambda h}\left(\int_t^{t+h}\lambda(t^*)dt^*\right)(\lambda h+o(h))+\sum_{k=2}^{\infty}\mathbb{P}(N_c(t+h)-N_c(t)=1,N(t+h)-N(t)=k)\\
=&\,\frac{1}{\lambda h}\left(\int_t^{t+h}\left(\lambda(t)+\lambda'(t)(t^*-t)+\lambda''(t)\frac{(t^*-t)^2}{2!}+\dots\right)dt^*\right)(\lambda h+o(h))+o(h)\\
&\text{(Expanding }\lambda(t^*)\text{ about }t\\
&\text{and see Remark 2 above to understand why the term in blue is }o(h))\\
=&\,\frac{1}{\lambda h}\left(\int_0^{h}\left(\lambda(t)+\lambda'(t)z+\lambda''(t)\frac{z^2}{2!}+\dots\right)dz\right)(\lambda h+o(h))+o(h)\\
=&\,\frac{1}{\lambda h}\left(\lambda(t)h+\lambda'(t)\frac{h^2}{2!}+\lambda''(t)\frac{h^3}{3!}+\dots\right)(\lambda h+o(h))+o(h)\\
=&\,\lambda(t)h+\frac{\lambda(t)o(h)}{\lambda}+\left(\frac{\lambda'(t)h}{2!\lambda}+\frac{\lambda''(t)h^2}{3!\lambda}+\dots\right)(\lambda h+o(h))+o(h)\\
&\text{(Verify that the term in red is }o(h))\\
=&\,\lambda(t)h+o(h)
\end{aligned}
$$

4. $\mathbb{P}(N_c(t+h)-N_c(t)\geq 2)=o(h)$
   Since $\mathbb{P}(N_c(t+h)-N_c(t)\geq 2)\leq\mathbb{P}(N(t+h)-N(t)\geq 2)\leq o(h)$, and $\mathbb{P}(N_c(t+h)-N(t)\geq 2)\geq 0$.
   Now apply sandwich theorem to show that $\mathbb{P}(N_c(t+h)-N_c(t)\geq 2)=o(h)$.

This completes the proof that $N_c(t)$ is a non-homogeneous Poisson process with rate $\lambda(t)$.

Compound Poisson process

Consider that packages arrive at a warehouse following a Poisson process $N(t)$ with rate λ. Suppose the weight of package i is given by, W_i. The W_i's are iid random variables. Now we are interested in understanding the total weight of packages received by any time t. $N(t)$ and W_i are independent. This is given by $W(t)$,

$$W(t) = \sum_{i=1}^{N(t)} W_i$$

Note that $W(t)$ is a random sum. Both the summand and the number of terms in the summand are random. $W(t)$ is a **Compound Poisson process**.

Exercises

1. Show that $\mathbb{E}[W(t)] = \lambda t \mathbb{E}[W_1]$.
2. Show that $\text{Var}[W(t)] = (\lambda t)^2 \mathbb{E}[W_1^2]$.
3. Suppose families migrate to an area with rate $\lambda = 2$ per week. The number of people in each family is independent. The probability of a family of size 1 is 1/6, 2 is 1/3, 3 is 1/3 and 4 is 1/6. Find the expected number of people migrating in 5 weeks.

References

1. Sheldon Ross, Introduction to Probability Models, Academic Press, 2024.
2. Rick Durrett, Essentials of Stochastic Processes, Springer, 1999.
3. Sheldon Ross, Stochastic Processes, Wiley, 2018.