ECE C143A/C243A - Neural signal processing and machine learning
Prof. Jonathan Kao
(with notes adapted from Prof. Byron Yu, CMU)

Poisson Processes

1 The exponential distribution

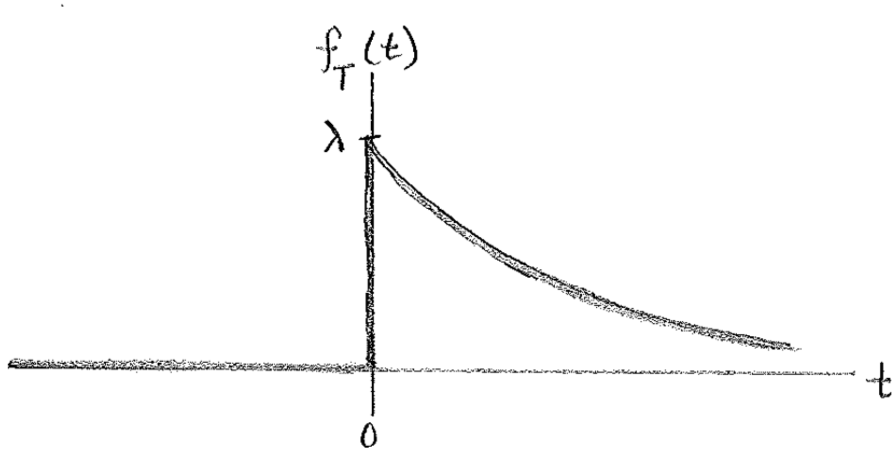
To understand and construct a Poisson process, we need a solid understanding of the exponential distribution and its properties.

1.1 Density function

A random variable T is said to be exponentially distributed with rate $\lambda > 0$ if its probability density function (pdf) is given by,

$$\boxed{f_T(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}} \tag{1}$$

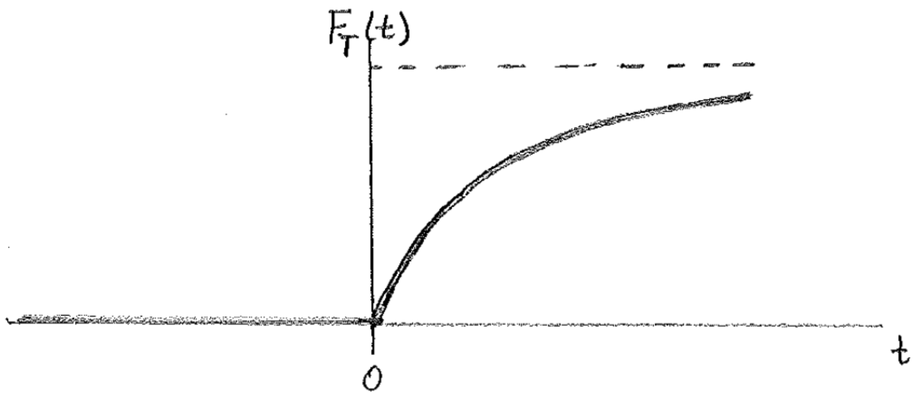
We denote this as $T \sim \exp(\lambda)$. The exponential pdf can be plotted as below:



Aside: The reason we care about the exponential distribution is that the time between two consecutive spikes (called the interspike interval, or ISI) can be modeled by an exponential distribution.

The cumulative distribution function (cdf) of the exponentially distributed random variable T is:

$$\boxed{F_T(t) = \begin{cases} 1 - e^{-\lambda t} & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}} \tag{2}$$



Aside: The relationship between the pdf and the cdf is given by the following relations:

$$\begin{aligned} f_T(t) &= \frac{dF_T(t)}{dt} \\ F_T(t) &= \int_{-\infty}^{t} f_T(t')dt' \end{aligned} \tag{3}$$

This is true for any distribution, not just the exponential. Why?

We can write the expression for a probability in terms of both the pdf and the cdf. In particular, note that with the cdf,

$$\begin{aligned} \mathbf{Pr}\ (t < T \leq t + \varepsilon) &= F_T(t+\varepsilon) - F_T(t) \\ &= \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon}\varepsilon \end{aligned} \tag{4}$$

Similarly, for the pdf and sufficiently small ε, we have that

$$\mathbf{Pr}\ (t < T \leq t + \varepsilon) \approx f_T(t)\varepsilon \tag{5}$$

By comparing expressions for the cdf and pdf, and setting $\varepsilon \to 0$, we arrive at:

$$\begin{aligned} f_T(t) &= \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon} \\ &= \frac{dF_T(t)}{dt} \end{aligned}$$

1.2 Mean and variance of the exponential distribution

We can calculate the mean of the random variable T as follows:

$$\begin{aligned} \mathbb{E}[T] &= \int t f_T(t) dt \\ &= \int_0^\infty t\lambda e^{-\lambda t} dt \end{aligned} \tag{6}$$

To perform this integral, we do integration by parts. In particular, we set $u = t$, $v = -e^{-\lambda t}$, $du = dt$, and $dv = \lambda e^{-\lambda t} dt$. Then,

$$\begin{aligned} \mathbb{E}[T] &= uv|_0^\infty - \int_0^\infty v du \\ &= -te^{-\lambda t}|_0^\infty + \int_0^\infty e^{-\lambda t} dt \\ &= 0 - 0 - \frac{1}{\lambda} e^{-\lambda t}\Big|_0^\infty \\ &= \frac{1}{\lambda} \end{aligned}$$

Thus, we have that:

$$\boxed{\mathbb{E}[T] = \frac{1}{\lambda}}$$

To calculate the variance of T, we need to calculate $\mathbb{E}[T^2]$. This can be done as follows:

$$\begin{aligned}\mathbb{E}[T^2] &= \int t^2 f_T(t)dt \\ &= \int_0^\infty t^2 \lambda e^{-\lambda t} dt\end{aligned}$$

Again, we perform integration by parts. This time, we let $u = t^2$, $v = -e^{-\lambda t}$, $du = 2tdt$, $dv = \lambda e^{-\lambda t}dt$. Then,

$$\begin{aligned}\mathbb{E}[T^2] &= uv|_0^\infty - \int_0^\infty vdu \\ &= -t^2 e^{-\lambda t}|_0^\infty + \int_0^\infty 2te^{-\lambda t}dt \\ &= 0 - 0 + \frac{2}{\lambda}\mathbb{E}[T] \\ &= \frac{2}{\lambda^2}\end{aligned}$$

Finally, we calculate **var**$[T] = \mathbb{E}[T^2] - (\mathbb{E}[T])^2$ to arrive at:

$$\boxed{\textbf{var}[T] = \frac{1}{\lambda^2}}$$

1.3 The exponential distribution is memoryless

In words, we can state the memoryless property informally as: say that the waiting time for a bus to arrive is exponentially distributed. If I've been waiting for t seconds, then the probability that I must wait s more seconds is the same as if I hadn't waited at all.

To codify this as a mathematical statement, we say that:

$$\boxed{\text{Memoryless property}: \quad \textbf{Pr}\,(T > t + s | T > t) = \textbf{Pr}\,(T > s)} \tag{7}$$

To show this holds for the exponential distribution, we calculate the conditional probability,

$$\begin{aligned}
\mathbf{Pr}\,(T > t + s | T > t) &= \frac{\mathbf{Pr}\,(T > t + s)}{\mathbf{Pr}\,(T > t)} \\
&= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} \\
&= e^{-\lambda s} \\
&= \mathbf{Pr}\,(T > s).
\end{aligned}$$

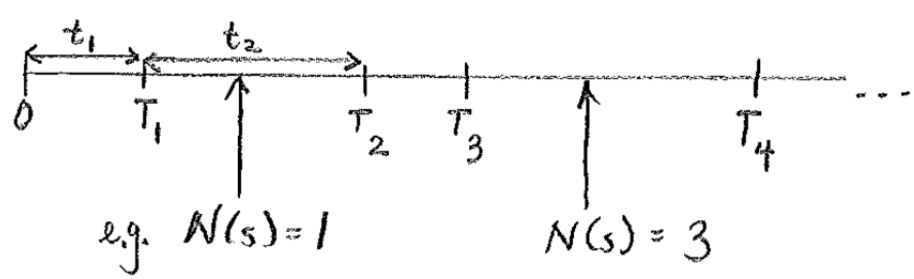
2 Defining the Poisson process

In this section, we will use the exponential distribution to construct a Poisson process.

2.1 Constructing a Poisson process

Let $t_1, t_2, \ldots$ be independent exponential random variables with parameter λ. Let $T_n = t_1 + t_2 + \cdots + t_n$ for $n \geq 1$ with $T_0 = 0$. Define $N(s) = \max\{n : T_n \leq s\}$.

$N(s)$ is a Poisson process.



If a Poisson process is used to model a spike train, then we have the following construction:

- t_n is the n^{th} interspike interval (ISI).
- T_n is the time at which the n^{th} spike occurs.

- $N(s)$ is the number of spikes by time s.
- λ is the neuron's firing rate.

2.2 $N(s)$ has a Poisson distribution with mean λs

Why is $N(s)$ called a Poisson process rather than an exponential process? Recall that we constructed $N(s)$ by defining exponentially distributed intervals. The reason is that the Poisson distribution comes into play as we look at the properties of $N(s)$.

$$\boxed{N(s) \text{ has a Poisson distribution with mean } \lambda s.}$$

To demonstrate this, we first recognize that

$$N(s) = n \iff T_n \leq s < T_{n+1}$$

That is, the n^{th} spike occurs before time s and the $(n+1)^{th}$ spike occurs after time s. Therefore,

$$\textbf{Pr}\,(N(s) = n) = \textbf{Pr}\,(T_{n+1} > s, T_n \leq s)$$

What is the intuition behind how we can simplify this expression? To build intuition, let's say that the times of the spikes were discrete. Using the chain rule,

$$\begin{aligned}
\textbf{Pr}\,(T_{n+1} > s, T_n \leq s) &= \textbf{Pr}\,(T_{n+1} > s | T_n \leq s)\textbf{Pr}\,(T_n \leq s) \\
&= \textbf{Pr}\,(T_{n+1} > s | T_n = s)\textbf{Pr}\,(T_n = s) \\
&\quad + \textbf{Pr}\,(T_{n+1} > s | T_n = s-1)\textbf{Pr}\,(T_n = s-1) \\
&\quad + \textbf{Pr}\,(T_{n+1} > s | T_n = s-2)\textbf{Pr}\,(T_n = s-2) \\
&\quad + \quad \vdots \\
&\quad + \textbf{Pr}\,(T_{n+1} > s | T_n = 0)\textbf{Pr}\,(T_n = 0) \\
&= \sum_{t=0}^{s} \textbf{Pr}\,(T_{n+1} > s | T_n = t)\textbf{Pr}\,(T_n = t)
\end{aligned}$$

Now with this intuition, we can return back to the continuous domain and replace the sum with an integral and $\textbf{Pr}\,(T_n = t)$ with the probability density function $f_{T_n}(t)$. This gives

$$\textbf{Pr}\,(T_{n+1} > s, T_n \leq s) = \int_0^s \textbf{Pr}\,(T_{n+1} > s | T_n = t) f_{T_n}(t) dt$$

and thus,

$$\begin{aligned}
\textbf{Pr}\,(N(s)=n) &= \textbf{Pr}\,(T_{n+1}>s, T_n \le s) \\
&= \int_0^s \textbf{Pr}\,(T_{n+1}>s|T_n=t) f_{T_n}(t)dt \\
&= \int_0^s \textbf{Pr}\,(t_{n+1}>s-t) f_{T_n}(t)dt
\end{aligned}$$

To simplify this expression, we need to find an expression for $f_{T_n}(t)$, the pdf of the time of the n^{th} spike.

Recall that $T_n = t_1 + t_2 + \cdots + t_n$, where $t_1, \ldots, t_n \sim \exp(\lambda)$ iid. Further, recall from your previous probability course (perhaps) that the distribution of a sum of independent random variables is the convolution of their pdf's. To easily convolve pdf's, we can take the Fourier transform and thus multiply rather than convolve. Concretely,

$$\begin{aligned}
\mathcal{F}(f_{T_n}) &= \prod_{i=1}^{n} \mathcal{F}(f_{t_i}) \\
&= \left(\mathcal{F}\left(\lambda e^{-\lambda t} u(t)\right)\right)^n \\
&= \left(\frac{\lambda}{\lambda + j\omega}\right)^n.
\end{aligned}$$

Taking $\mathcal{F}^{-1}$ of both sides,

$$\begin{aligned}
f_{T_n}(t) &= \frac{\lambda^n}{(n-1)!} t^{n-1} e^{-\lambda t} u(t) \\
&= \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \qquad \text{for } t \ge 0
\end{aligned}$$

This is called the *Erlang distribution* which is a special case of the *gamma distribution*. We will see this distribution again when we try to model the refractory period.

Back to calculating $\textbf{Pr}\,(N(s)=n)$, we proceed via,

$$\begin{aligned}
\textbf{Pr}\left(N(s)=n\right) &= \int_0^s e^{-\lambda(s-t)}\lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}dt \\
&= \frac{\lambda^n}{(n-1)!}e^{-\lambda s}\int_0^s t^{n-1}dt \\
&= \frac{\lambda^n}{(n-1)!}e^{-\lambda s}\left[\frac{t^n}{n}\right]_0^s \\
&= e^{-\lambda s}\frac{(\lambda s)^n}{n!}
\end{aligned}$$

and therefore, $N(s) \sim \text{Poisson}(\lambda s)$.

Let's calculate the mean and variance of $N(s)$.

$$\begin{aligned}
\mathbb{E}[N(s)] &= \sum_{n=0}^{\infty} n\textbf{Pr}\left(N(s)=n\right) \\
&= \sum_{n=1}^{\infty} ne^{-\lambda s}\frac{(\lambda s)^n}{n!} \\
&= \lambda s\sum_{n=1}^{\infty} e^{-\lambda s}\frac{(\lambda s)^{n-1}}{(n-1)!} \\
&= \lambda s.
\end{aligned}$$

To calculate the variance of $N(s)$, we note that

$$\begin{aligned}
\textbf{var}[N(s)] &= \mathbb{E}[N(s)^2] - (\mathbb{E}[N(s)])^2 \\
&= \mathbb{E}[N(s)(N(s)-1)] + E[N(s)] - (E[N(s)])^2,
\end{aligned}$$

which means we need to calculate $\mathbb{E}[N(s)(N(s)-1)]$.

$$\begin{aligned}
\mathbb{E}[N(s)(N(s)-1)] &= \sum_{n=0}^{\infty} n(n-1)\textbf{Pr}\left(N(s)=n\right) \\
&= \sum_{n=2}^{\infty} n(n-1)e^{-\lambda s}\frac{(\lambda s)^n}{n!} \\
&= (\lambda s)^2,
\end{aligned}$$

and thus,

$$\begin{aligned}\textbf{var}(N(s)) &= (\lambda s)^2 + \lambda s - (\lambda s)^2 \\ &= \lambda s.\end{aligned}$$

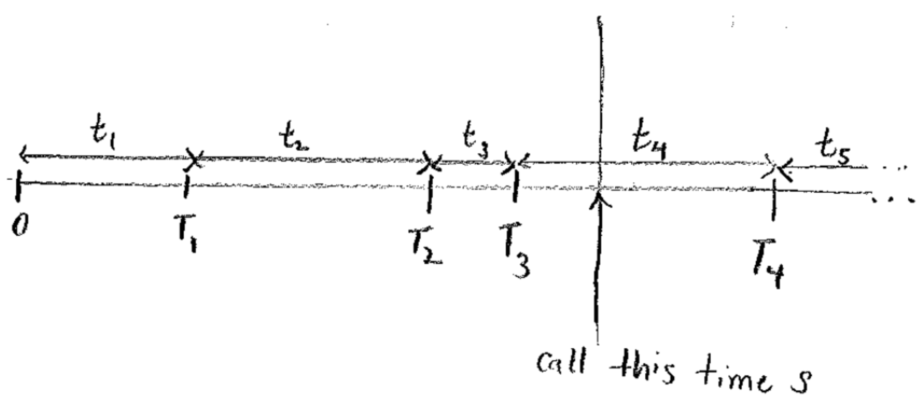
2.3 Increments of a Poisson process are independent of the past

A second property of the Poisson process is that the number of spikes in any increment is independent of the past. Mathematically,

$$\boxed{N(t+s) - N(s), t \geq 0, \text{ is a rate } \lambda \text{ Poisson process and independent of } N(r) \text{ for } 0 \leq r \leq s.}$$

In words, this means that if you look forward from any time s, that is itself a Poisson process independent of anything that's already happened.

We won't prove this formally but the following provides intuition:



Looking forward from time s, the time until the first spike (at T_4) is distributed $\exp(\lambda)$ and independent of anything that came before it, by the memoryless property of the exponential. Subsequent ISI's $(t_5, t_6, \ldots)$ are distributed $\exp(\lambda)$ and independent of anything before time s.

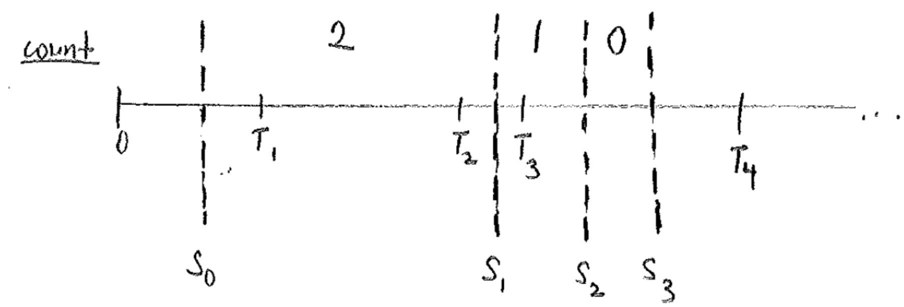
2.4 $N(t)$ has independent increments

A third property of the Poisson process is that

$$\boxed{N(t) \text{ has independent increments.}}$$

Mathematically, if $s_0 < s_1 < \cdots < s_n$, then, $N(s_1)-N(s_0), N(s_2)-N(s_1), \ldots, N(s_n)-N(s_{n-1})$ are independent.

In other words, if you take spike counts in non-overlapping windows, the spike counts are independent.



2.5 Summary of the Poisson process properties

If $\{N(s), s \geq 0\}$ is a Poisson process with rate λ, then

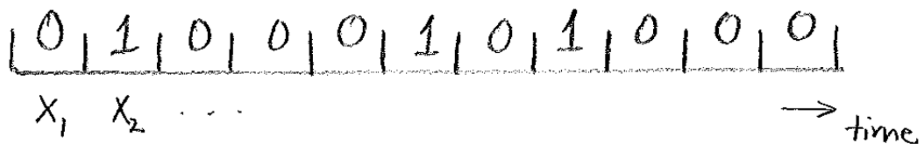
1. $N(0) = 0$
2. $N(t+s) - N(s) \sim$ Poisson(λt)
3. $N(s)$ has independent increments

The converse is also true. That is, if these three properties hold, then $\{N(s), s \geq 0\}$ is a Poisson process.

2.6 Another view of the Poisson process

So far, we have derived the Poisson process using iid exponential ISI's. Another very useful way of thinking about the Poisson process is using the Bernoulli process.

The Poisson process is the continuous-time limit of the Bernoulli process, which is defined in discrete time.



Let n be the number of discrete time steps. Let p be the probability of spiking at each time step. In the Bernoulli process, at each time step, we flip a coin and decide whether the neuron spikes (1) or not (0). The coin flips are independent of each other.

At the i^{th} time step,

$$X_i \sim \text{Bernoulli}(p) \qquad \text{iid}$$

with,

$$X_i = \begin{cases} 1 & \text{w.p. } p \\ 0 & \text{w.p. } 1-p \end{cases} \tag{8}$$

Now, let S_n be the number of spikes up to and including the n^{th} time step.

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

As a sum of Bernoulli random variables, $S_n \sim \text{Binomial}(n, p)$. Therefore,

$$\mathbf{Pr}\,(S_n = k) = \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{(n-k)},$$

and $\mathbb{E}[S_n] = np$. Thus, we expect to see np spikes in n time steps.

As $n \to \infty$ and $p \to \infty$, the Bernoulli process becomes the Poisson process, where $np = \lambda s$. While we won't prove this, a key point is that the Bernoulli process provides an intuitive way to think about the Poisson process.

Another way to see this is to consider the probability that a Poisson process gives a spike in a small window of duration δ. The number of spikes in this window is $\sim$ Poisson$(\lambda\delta)$. Then,

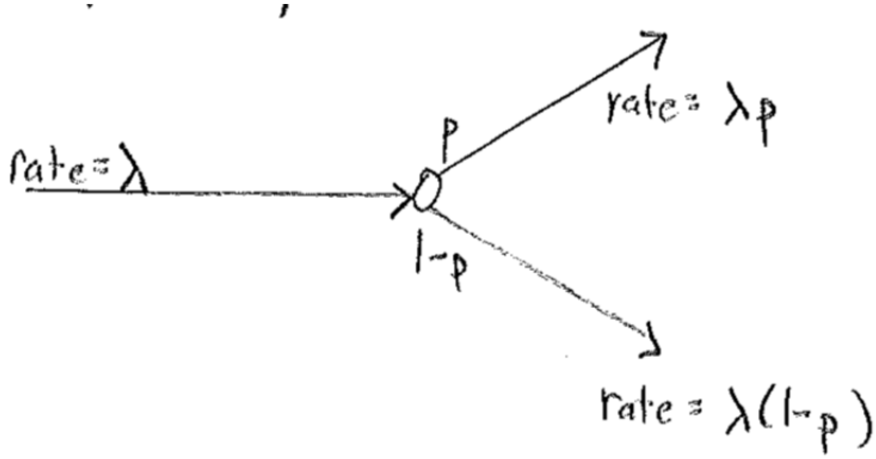
- **Pr** $(0$ spikes in $[t, t+\delta]) = e^{-\lambda\delta} = 1 - \lambda\delta + \mathcal{O}(\delta)^2$
- **Pr** $(1$ spike in $[t, t+\delta]) = \lambda\delta e^{-\lambda\delta} = \lambda\delta - \mathcal{O}(\delta)^2$
- **Pr** $(2$ spikes in $[t, t+\delta]) = \mathcal{O}(\delta)^2$

If δ is small, then $\mathcal{O}(\delta^2)$ terms go to 0. Hence, whether or not the neuron spikes in this window can be determined with a coin flip, where the probability of a spike is $\lambda\delta$.

3 Thinning and Superposition

3.1 Thinning

Suppose $N(s)$ is a Poisson process with rate λ. Each time a spike occurs, a coin is flipped. If the coin comes up heads (w.p. p) the spike is assigned to output stream 1. Else, the spike is assigned to output stream 2. The two output streams are each an independent Poisson process with rates λp and $\lambda(1-p)$ respectively.



3.2 Superposition

Suppose $N_1(s)$ and $N_2(s)$ are independent Poisson processes with rates λ_1 and λ_2 respectively. Then $N_1(s) + N_2(s)$ is a Poisson process with rate $\lambda_1 + \lambda_2$.

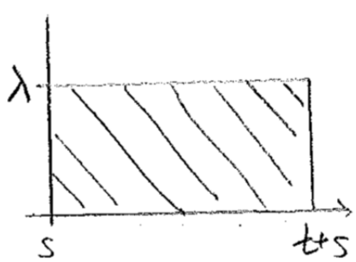
4 Inhomogeneous Poisson process

So far, we've considered the homogeneous Poisson process whose rate does not change with time. However, the firing rates of neurons typically *do* change with time. To model the time-dependent activity of neurons, we need a non-stationary process, such as the inhomogeneous Poisson process.

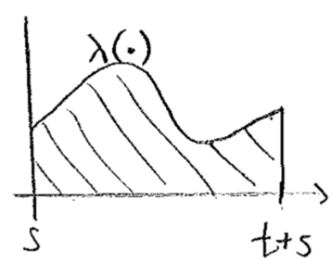
The inhomogeneous Poisson process is defined as follows: $\{N(s), s \geq 0\}$ is an inhomogeneous Poisson process with rate $\lambda(r)$ if

1. $N(0) = 0$.
2. $N(t+s) - N(s) \sim \text{Poisson}\left(\int_s^{t+s} \lambda(r)dr\right)$
3. $N(s)$ has independent increments.

Comparing with a homogeneous Poisson process, the only difference is that the Poisson mean is now $\int_s^{t+s} \lambda(r)dr$ rather than λt.



Homogeneous case



Inhomogeneous case

Note that if $\lambda(r)$ is flat, then the two definitions are equivalent.

4.1 ISIs of an inhomogeneous Poisson process are NOT independent

What is the interarrival distribution of the Poisson process? In general, this is not a simple expression, and was derived by Yakovlev and colleagues (Yakovlev et al., arXiv 2005, see `https://arxiv.org/pdf/cond-mat/0507657.pdf`). Even though this expression is complicated, we can still show that the interarrival times are neither exponential or independent.

To show, this, let $\mu(t) = \int_0^t \lambda(r)dr$. Then, we can at least get the distribution of the first ISI.

$$\begin{aligned}
\textbf{Pr}\,(t_1 > t) &= \textbf{Pr}\,(N(t) = 0) \\
&= e^{-\int_0^t \lambda(r)dr} \\
&= e^{-\mu(t)}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
f_{t_i}(t) &= -\frac{d}{dt}\textbf{Pr}\,(t_1 > t) \\
&= \lambda(t)e^{-\mu(t)}.
\end{aligned}$$

From this, we can see a resemblance to the exponential distribution, but already we see that it doesn't take on the exact form. And thus we can say that the distribution of the ISIs is not exponential.

Now, are the ISI's independent? To evaluate this, we look forward in time from T_1.

$$\begin{aligned}
\textbf{Pr}\,(t_2 > s | t_1 = t) &= \textbf{Pr}\,(N(t+s) - N(t) = 0 | t_1 = t) \\
&= \textbf{Pr}\,(N(t+s) - N(t) = 0 | N(t) = 1, N(r) = 0) \text{ for } 0 \le r < t \\
&= \textbf{Pr}\,(N(t+s) - N(t) = 0) \quad \text{(by independent incr.)} \\
&= e^{-\int_t^{t+s} \lambda(r)dr} \\
&= e^{-(\mu(s+t) - \mu(t))}
\end{aligned}$$

Note that in the 2nd to 3rd line, we used the fact that the Poisson process has independent increments, and thus the distribution of $N(t+s) - N(s)$ is independent of anything that happened before and up to time t. From this, we can calculate the conditional pdf,

$$
\begin{aligned}
f_{t_2|t_1}(s) &= -\frac{d}{ds}\mathbf{Pr}\left(t_2 > s | t_1 = t\right) \\
&= \lambda(s+t)e^{-(\mu(s+t)-\mu(t))}
\end{aligned}
$$

Since t_2 depends on the value t_1 takes on, the ISI's are not independent.

4.2 Joint distribution of the ISIs

With these results, we can calculate the joint distribution of the ISIs:

$$
\begin{aligned}
f_{t_1,t_2}(t,s) &= f_{t_1}(t)f_{t_2|t_1}(s|t) \\
&= \lambda(t)\lambda(t+s)e^{-\mu(s+t)}
\end{aligned}
$$

If we change variables from ISIs to spike times, we arrive at:

$$f_{T_1,T_2}(\nu_1,\nu_2) = \lambda(\nu_1)\lambda(\nu_2)e^{-\mu(\nu_2)}$$

and more generally,

$$f_{T_1,\ldots,T_N}(\nu_1,\ldots,\nu_n) = \lambda(\nu_1)\cdots\lambda(\nu_n)e^{-\mu(\nu_n)} \tag{9}$$

In words, this equation tells us that the spike train probability density is equal to sampling $\lambda(r)$ at the spike times, and then multiplying by the exponentiated integral of the rate up until the last observed spike. An important observation is that this density is a function of the times of each spike.

In contrast, what does this density equal for a homogeneous Poisson process? To write this expression for a homogeneous Poisson process, we substitute $\lambda(r) = \lambda_0$ for all r, where λ_0 is the rate of the homogeneous Poisson process. In this case,

$$f_{T_1,\ldots,T_n}(\nu_1,\ldots,\nu_n) = \lambda_0^n e^{-\lambda_0\nu_n} \tag{10}$$

Note that this does not depend on $\nu_1,\ldots,\nu_{n-1}$. The probability of a spike train depends only on the number of spikes n and the time of the last spike, ν_n. Note, it would have also been possible to arrive at equation (8) by multiplying exponential distributions, since the ISIs are iid, i.e.,

$$
\begin{aligned}
f_{t_1,\ldots,t_N}(u_1,\ldots,u_n) &= \prod_{i=1}^{n}\lambda_0 e^{-\lambda_0 u_i} &(11)\\
&= \lambda_0^n e^{-\lambda\sum_{i=1}^{n}u_i} &(12)
\end{aligned}
$$

where $\nu_n = \sum_{i=1}^{n} u_i$.

The key take-home point here is a contrast in the homogeneous and inhomogeneous Poisson processes. For a homogeneous Poisson process, the probabillity of a given spike train is only a function of the window over which you look. Further, the probability of seeing a specific sequence of n spikes is the same across all sequences (i.e., you could put these n spikes anywhere in time). (Does this make intuitive sense? It may be helpful to think about the Bernoulli process.) On the other hand, because the rate of the inhomogeneous Poisson process is now time-dependent, the spike train density for a sequence with n spikes depends exactly on where each spike is. Not all sequences of an inhomogeneous Poisson process are equally probable.

5 Generating Poisson processes

5.1 Homogeneous Poisson process with rate λ

We'd like to generate a Poisson process of length $\mathcal{T}$:

Method 1:

- Generate iid exponential random variables $t_1, t_2, \ldots$ with parameter λ. (In `MATLAB`, use `exprnd`).
- The spike times are $T_n = \sum_{i=1}^{n} t_i$.
- If $T_n > \mathcal{T}$, stop.

Method 2:

- Draw $N \sim$ Poisson($\lambda\mathcal{T}$), the number of spikes on the interval $[0, \mathcal{T}]$. (In `MATLAB`, use `poissrnd`).
- Draw $T_1, \ldots, T_N \sim$ Uniform($[0, \mathcal{T}]$). (In `MATLAB`, use `rand`). The $T_1, \ldots, T_N$ are the spike times.

Aside: Why does Method 2 work? The intuition is that a spike should not be more likely to occur at one time compare to another time. More formally, Method 2 is based on the theorem that if we condition on $N(\mathcal{T}) = N$, then the set of spike times $\{T_1, \ldots, T_N\}$ has the same distribution as $\{U_1, \ldots, U_N\}$ where $U_1, \ldots, U_N \sim$ Uniform($[0, \mathcal{T}]$) iid.

5.2 Inhomogeneous Poisson process with rate $\lambda(t)$

- Let $\lambda_{\text{max}} = \max_t \lambda(t)$. Genereate a homogeneous Poisson process with rate λ_{max}.
- For $n = 1, \ldots, N$,
  - Draw $U \sim$ Uniform$([0, 1])$. If $U > \lambda(T_n)/\lambda_{\text{max}}$, reject the spike at T_n. Else, retain the spike at T_n.

The spikes that are retained at the end of this procedure is an inhomogeneous Poisson process with rate $\lambda(t)$. (We won't prove this.)