Poisson Processes

18-698 / 42-632

Neural Signal Processing
Prof. Byron Yu

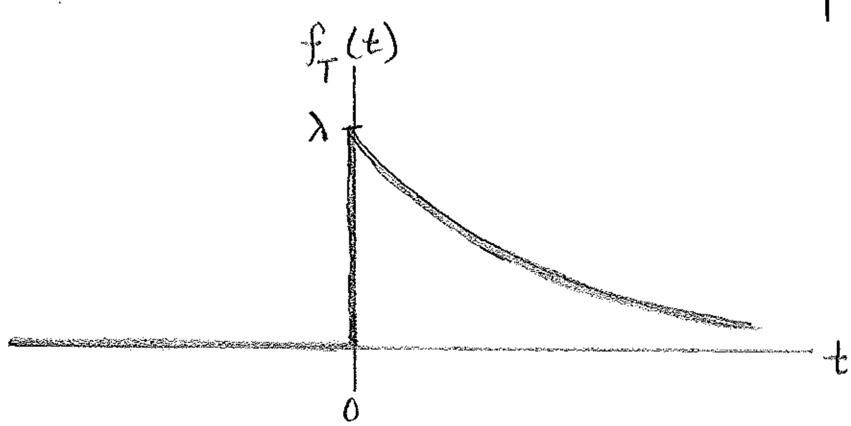
A) Exponential Distribution

To prepare for our discussion of Poisson processes, we will need to go through some properties of the exponential distribution first.

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

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

As a shorthand, we can write $T \sim \exp(\lambda)$.



[A peek ahead:
The time between two consecutive spikes (a.k.a. the interspike interval, or ISI) can be modeled by an exponential distribution.]

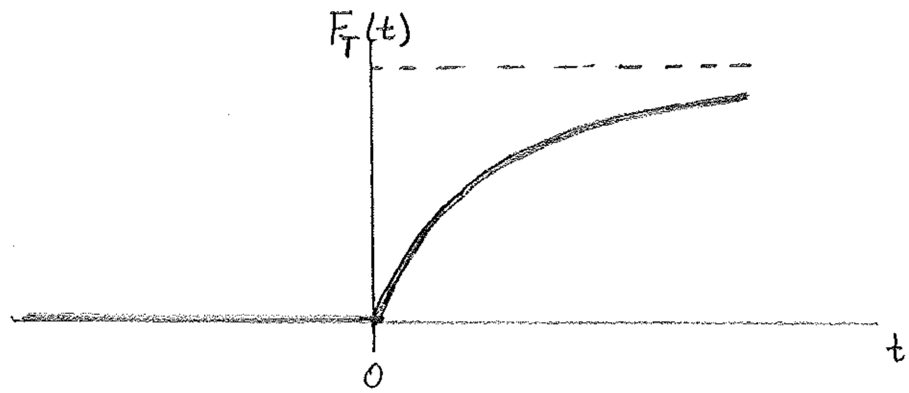
Alternatively, we can describe T in terms of its cumulative distribution function (CDF)

exponential

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

$F \rightarrow$ cdf

$f \rightarrow$ pdf



Note that the PDF and CDF are related in the following way:

pdf is derivative of cdf w/ respect to time

$$f_T(t) = \frac{d F_T(t)}{dt} \qquad F_T(t) = \int_{-\infty}^{t} f_T(t')\, dt' \quad (3)$$

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

Using the CDF,

$$P(t < T \leq t + \varepsilon) = F_T(t+\varepsilon) - F_T(t) \quad (4)$$

$$= \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon} \cdot \varepsilon$$

Using the PDF,

$$P(t < T \leq t+\varepsilon) \approx f_T(t) \cdot \varepsilon \qquad (5)$$

Comparing (4) and (5) as $\varepsilon \to 0$,

$$f_T(t) = \frac{F_T(t+\varepsilon) - F_T(t)}{\varepsilon} = \frac{d\,F_T(t)}{dt}$$

//

A.1) Mean and variance of the exponential

$$E[T] = \int t \, f_T(t)\, dt$$

$$= \int_0^\infty t \cdot \lambda e^{-\lambda t} \, dt \qquad (6)$$

Integrating by parts, let $u = t$ $\quad dv = \lambda e^{-\lambda t} dt$

$du = dt \quad v = -e^{-\lambda t}$

$$E[T] = uv \Big|_0^\infty - \int_0^\infty v \, du$$

$$= -t e^{-\lambda t} \Big|_0^\infty + \int_0^\infty e^{-\lambda t} \, dt$$

$$= 0 - 0 + \left[-\frac{1}{\lambda} e^{-\lambda t} \right]_0^\infty$$

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

exponential

$$E[T^2] = \int t^2 f_T(t)\,dt$$
$$= \int_0^\infty t^2 \cdot \lambda e^{-\lambda t}\,dt$$

Integrating by parts, let $u = t^2$ $\quad dv = \lambda e^{-\lambda t}\,dt$
$\quad du = 2t\,dt \quad v = -e^{-\lambda t}$

$$E[T^2] = uv\Big|_0^\infty - \int_0^\infty v\,du$$
$$= -t^2 e^{-\lambda t}\Big|_0^\infty + \int_0^\infty 2t \cdot e^{-\lambda t}\,dt$$
$$= 0 - 0 + \frac{2}{\lambda}\left(\int_0^\infty \lambda t\, e^{-\lambda t}\,dt\right)$$
we saw this in (6)!
$$= \frac{2}{\lambda^2}$$

$$\mathrm{Var}(T) = E[T^2] - (E[T])^2 = \boxed{\frac{1}{\lambda^2}}$$ exponential

A.2) Memoryless property of exponential

In words:

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.

With math:

$$P(T > t+s \mid T>t) = P(T > s) \qquad (7)$$

To show this,

$$P(T> t+s \mid T> t) = \frac{P(T > t+s)}{P(T>t)} \quad \text{from (2)}$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}}$$

$$= e^{-\lambda s}$$

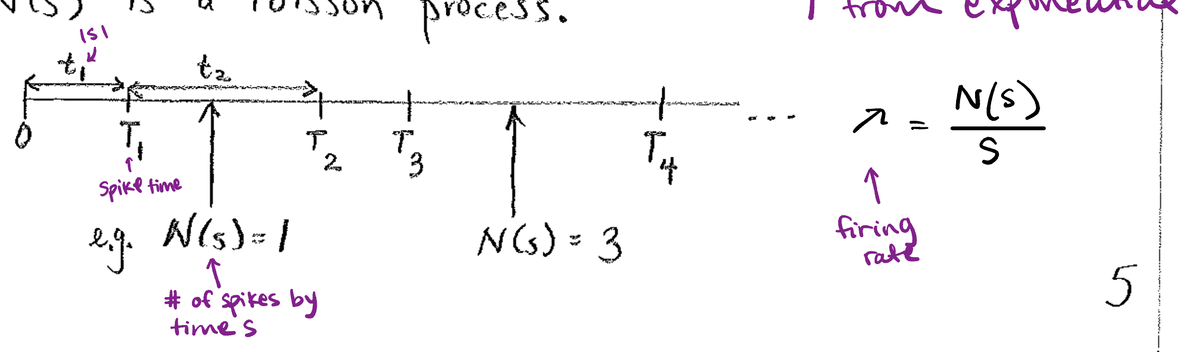
$$= P(T>s)$$

B) Defining the Poisson process

B.1) Constructing a Poisson process

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

$N(s)$ is a Poisson process. T from exponential



If a Poisson process is used to model a spike train, then:

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. Poisson

λ is the neuron's firing rate

B.2) Properties of the Poisson process

Why is $N(s)$ called a Poisson process rather than an exponential process?

$\boxed{\text{Property 1}}$: $N(s)$ has a Poisson distribution with mean λs.

Let's show this.

First, recognize that $N(s)=n$ iff: $T_n \le s < T_{n+1}$.
In other words, the n^{th} spike occurs before time s and the $(n+1)^{th}$ spike occurs after time s.

1:1 relationship
$$P(N(s)=n) = \int_0^s P(\overset{\text{after}}{T_{n+1} > s} \mid \overset{\text{at some time } t}{T_n = t}) \overset{\text{pdf}}{f_{T_n}}(\underset{\text{time of } n^{th} \text{ spike}}{t})\, dt$$

$$= \int_0^s P(t_{n+1} > s-t)\, \overset{\text{pdf}}{f_{T_n}}(\underset{\text{time of } n^{th} \text{ spike}}{t})\, dt$$

$f_{T_n}(t)$ is the PDF of the time of the n^{th} spike

As defined previously, $T_n = t_1 + t_2 + \dots + t_n$, where $t_1, \dots, t_n \sim \exp(\lambda)$ iid.

Recall from your previous probability course (perhaps) that summing independent random variables implies convolving their PDF's.

If we take Fourier transforms of the PDF's, then we can multiply rather than convolve.

$$\mathcal{F}\{f_{T_n}\} = \prod_{i=1}^{n} \mathcal{F}\{f_{t_i}\}$$

Fourier

$$= \left[\mathcal{F}\{\lambda e^{-\lambda t} u(t)\}\right]^n$$

$$= \left(\frac{\lambda}{\lambda + j\omega}\right)^n$$

Table of Fourier transforms

$$e^{-at} u(t) \xrightarrow{\mathcal{F}} \frac{1}{a+j\omega}$$

$$t^n e^{-at} u(t) \xrightarrow{\mathcal{F}} \frac{n!}{(a+j\omega)^{n+1}}$$

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

$$f_{T_n}(t) = \frac{\lambda^n}{(n-1)!} \cdot t^{n-1} e^{-\lambda t} u(t)$$

$$f_{T_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \quad \text{for } t \geq 0$$

pdf of time of n^{th} spike

This is called the Erlang distribution, which is a special case of the gamma distribution.

This will appear again when we try to model refractory periods.

$$P(N(s)=n) = \int_0^s e^{-\lambda(s-t)} \cdot \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}\, dt$$

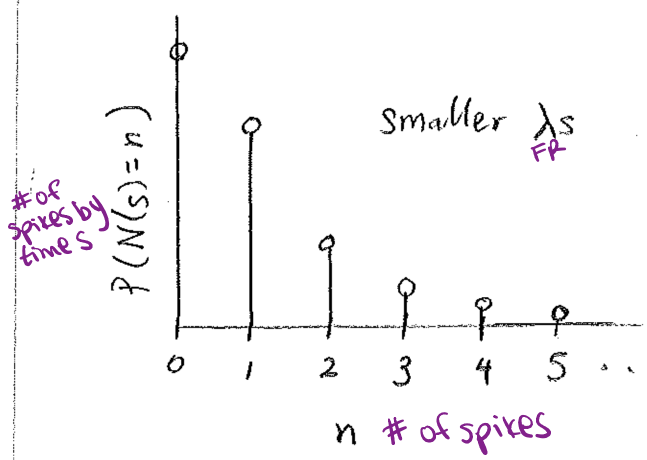
$$= \frac{\lambda^n}{(n-1)!} e^{-\lambda s} \int_0^s t^{n-1}\, dt$$

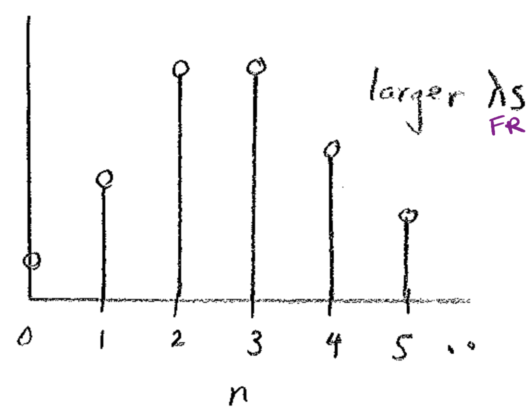
$$= \quad " \quad \left[\frac{t^n}{n}\right]_0^s$$

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

$$\Rightarrow N(s) \sim \text{Poisson}(\lambda s)$$

- What does a Poisson distribution look like?





- What is its mean and variance?

$$E[N(s)] = \sum_{n=0}^{\infty} \overset{\#\text{ of spikes}}{n} \cdot \overset{\#\text{ of spikes by time } s}{P(N(s)=n)}$$

mean of # spikes by time s

n=0 term doesn't contribute to sum

$$= \sum_{n=1}^{\infty} n \cdot e^{-\lambda s} \frac{(\lambda s)^n}{n!}$$

$$= \lambda s \underbrace{\sum_{n=1}^{\infty} e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!}}_{1}$$

summing over entire Poisson distribution

Poisson

units: spikes/sec

$$E[N(s)] = \boxed{\lambda s}$$

units: sec

mean # of spikes by time s

$\frac{\text{spikes}}{\cancel{s}} \cdot \cancel{s}$

To find the variance, we use a similar trick.

$$E\left[N(s)(N(s)-1)\right] = \sum_{n=0}^{\infty} n(n-1)\, P(N(s)=n)$$

n=0 and n=1 terms don't contribute to sum

$$= \sum_{n=2}^{\infty} n(n-1)\, e^{-\lambda s} \frac{(\lambda s)^n}{n!}$$

$$= (\lambda s)^2 \underbrace{\sum_{n=2}^{\infty} e^{-\lambda s} \frac{(\lambda s)^{n-2}}{(n-2)!}}_{1}$$

summing over entire Poisson distribution

$$= (\lambda s)^2$$

$$\operatorname{Var}(N(s)) = E[\underbrace{N(s)^2}_{\text{don't know}}] - (\underbrace{E[N(s)]}_{\text{we know this}})^2$$

$$= \underbrace{\boxed{E\left[N(s)(N(s)-1)\right]}}_{(\lambda s)^2} + \underbrace{E[N(s)]}_{\lambda s} - \underbrace{(E[N(s)])^2}_{(\lambda s)^2}$$

$$= \cancel{(\lambda s)^2} + \lambda s - \cancel{(\lambda s)^2}$$

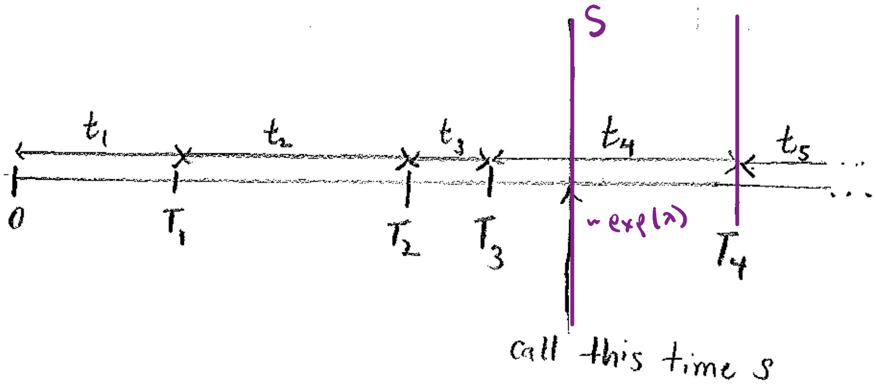
Poisson

$$\operatorname{Var}[N(s)] = \boxed{\lambda s}$$

Property 2: $N(t+s) - N(s)$, $t \geq 0$ is a rate λ Poisson process and independent of $N(r)$, $0 \leq r \leq s$.

In other words, 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 the intuition:



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

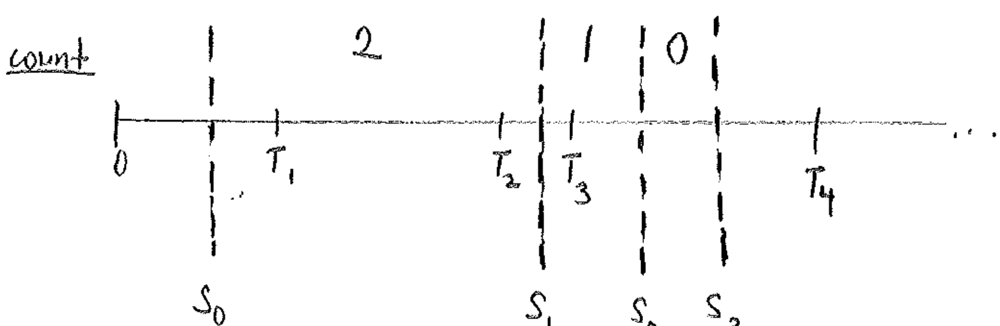
$\boxed{\text{Property 3}}$: $N(t)$ has independent increments.

If $s_0 < s_1 < \ldots < s_n$, 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.



To summarize,

Deriving Poisson process using iid exponential ISI's

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

(i) $N(0) = 0$

(ii) $N(t+s) - N(s) \sim \text{Poisson}(\lambda t)$

\# of spikes b/t t & s

(iii) $N(s)$ has independent increments.

Conversely, if (i), (ii), and (iii) hold, then $\{N(s), s \geq 0\}$ is a Poisson process.

B.3) Another view of the Poisson process

So far, we have derived the Poisson process using i.i.d. 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.

Poisson $\longrightarrow$ Continuous time

Bernoulli $\longrightarrow$ Discrete time

Bernoulli process

0	1	0	0	0	1	0	1	0	0	0

$X_1 \quad X_2 \quad \dots \qquad \longrightarrow$ time

n is the number of discrete time steps

p is the probability of spiking at each time step.

At each timestep, flip a coin to decide whether the neuron spikes (1) or not (0). The coin flips are independent of each other.

At the i^{th} time step,

$$X_i \sim \text{Bernoulli}(p) \quad \text{i.i.d.}$$

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

Let S_n be the number of spikes up to and including the n^{th} timestep.

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

spike count

$S_n \sim \text{Binomial}(n, p)$

$$P(S_n = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$E[S_n] = \underset{\text{prob of spiking in window}}{\overset{\text{\# of time steps}}{n\,p}}$$

$\Rightarrow$ We expect to see np spikes in n time steps

$n \to \infty$ flipping coin more & more times

$p \to 0$ becoming less likely to get heads

As $n \to \infty$ and $p \to 0$, the Bernoulli process becomes the Poisson process, where

$$np = \lambda s$$

np: mean spike count for Bernoulli process in n time steps.

λs: mean spike count for Poisson process in window of duration s.

We won't prove this here,

The key is that the Bernoulli process provides an intuitive way to think about the Poisson process

---

We can also go in the other direction and ask what is the probability that a Poisson process gives a spike in a small time window of duration δ.

The number of spikes in this window is $\sim \text{Poisson}(\lambda \delta)$.

neuron FR

duration of time window

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

$P(0 \text{ spikes in } [t, t+\delta]) = e^{-\lambda \delta} \frac{(\lambda s)^0}{0!} = \boxed{1 - \lambda\delta} + O(\delta)^2$

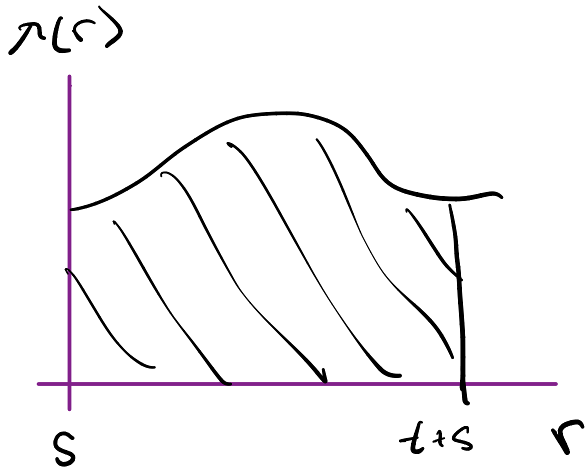
$P(1 \text{ spike in } [t, t+\delta]) = e^{-\lambda\delta} \cdot \frac{\lambda\delta}{1!} = \boxed{\lambda\delta} + O(\delta^2)$

"order of"

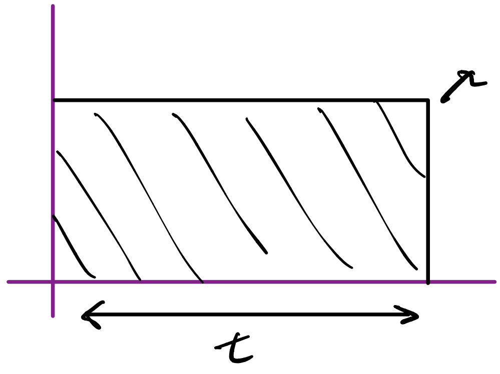
$P(>1 \text{ spike in } [t, t+\delta]) = O(\delta^2)$

If δ is small, $O(\delta^2)$ terms $\to 0$.

Whether or not the neuron spikes in this window can be determined with a coin flip, where the probability of a spike is $\lambda\delta$.



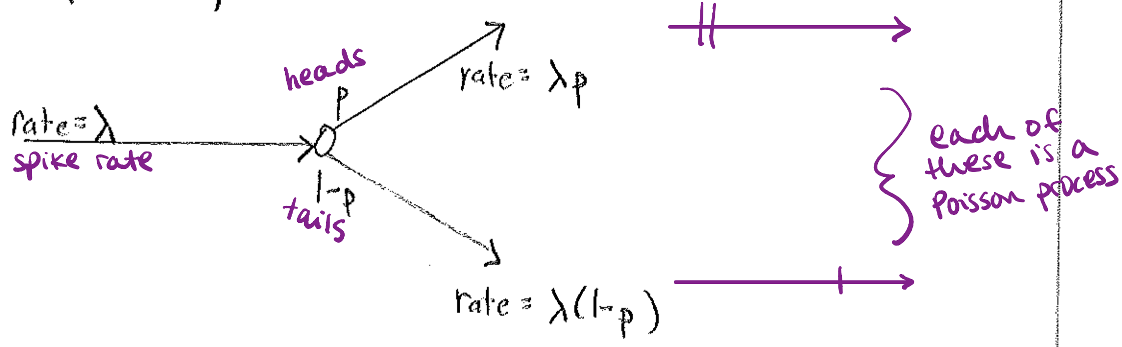
Inhomogenous



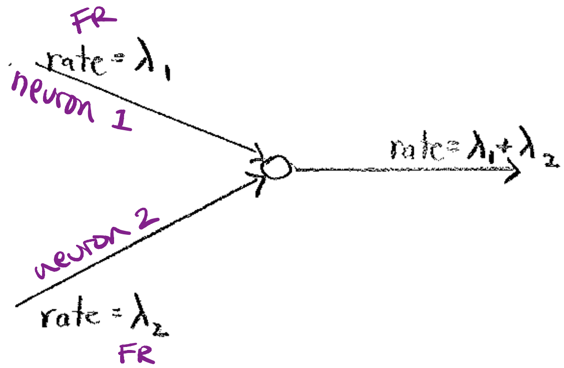
Homogenous

C) Thinning and Superposition

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.



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 $\lambda_1+\lambda_2$.



D) 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.

Definition

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

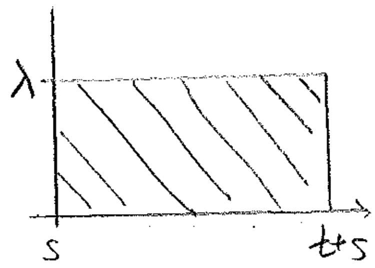
\# of spikes by time s

(i) $N(0) = 0$

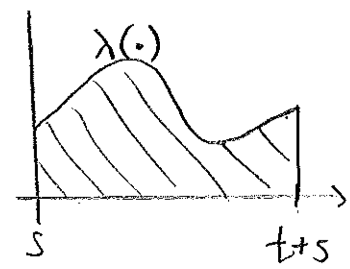
(ii) $N(t+s) - N(s) \sim \text{Poisson}\left(\int_s^{t+s} \lambda(r)\,dr \right)$

(iii) $N(s)$ has independent increments.

Comparing with the definition of a homogeneous Poisson process on p.11, the only difference is that the Poisson mean is now $\int_s^{t+s} \lambda(r)\,dr$ rather than λt.



Homogeneous case

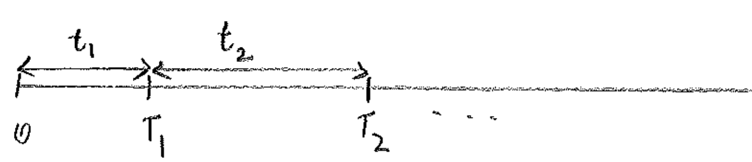


Inhomogeneous case

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

For an inhomogeneous Poisson process, the ISI's are no longer exponentially distributed or independent.

Let's show this.



Let $\mu(t) = \int_{0\ t}^{t+s} \lambda(r)\, dr$.

$$P(t_1 > t) = P(N(t) = 0) = e^{-\int_0^t \lambda(r)dr} = e^{-\mu(t)}$$

$$f_{t_1}(t) = -\frac{d}{dt} P(t_1 > t) = \lambda(t)\, e^{-\mu(t)} \Big\}$$ not exponential!

Now, look forward in time from T_1.

$$P(t_2 > s \mid t_1 = t) = P(N(t+s) - N(t) = 0)$$

Probability of # of spikes b/t t and t+s is 0

time of 1st spike

$$= e^{-\int_t^{t+s} \lambda(r)\, dr}$$

$$= e^{-(\mu(s+t) - \mu(t))}$$

$$f_{t_2|t_1}(s) = -\frac{d}{ds} P(t_2 > s \mid t_1 = t) = \lambda(s+t)\, e^{-(\mu(s+t) - \mu(t))}$$

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

The joint distribution of ISI's is

$$f_{t_1, t_2}(t, s) = f_{t_2|t_1}(s) \cdot f_{t_1}(t)$$

$$= \lambda(t)\lambda(t+s)\, e^{-\mu(s+t)}$$

Changing variables from ISI's to spike times,

$$f_{T_1, T_2}(v_1, v_2) = \lambda(v_1)\lambda(v_2)\, e^{-\mu(v_2)}$$

For more than 2 spikes,

Spike train probability density

$$\boxed{f_{T_1, \ldots, T_n}(v_1, \ldots, v_n) = \lambda(v_1) \ldots \lambda(v_n)\, e^{-\mu(v_n)}} \quad (7)$$

What does the spike train probability density (7) reduce down to for a homogeneous Poisson process?

For a homogeneous Poisson process, $\lambda(r) = \lambda_0 \; \forall r$.

$$f_{T_1,\ldots,T_n}(v_1,\ldots,v_n) = \lambda_0^n \, e^{-\lambda_0 v_n} \qquad (8)$$

Note that this does not depend on $v_1,\ldots,v_{n-1}$.

The probability of a spike train depends only on the number of spikes n and the time of the last spike v_n.

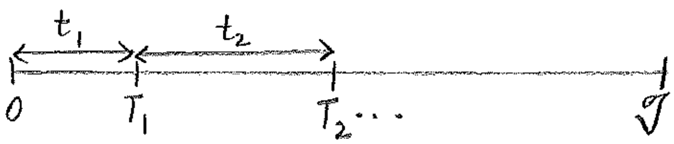
Equation (8) could also have been obtained by multiplying exponential distributions, since ISI's are i.i.d.

$$f_{t_1,\ldots,t_n}(u_1,\ldots,u_n) = \prod_{i=1}^{n} \lambda_0 e^{-\lambda_0 u_i}$$

$$= \lambda_0^n \, e^{-\lambda_0\left(\sum_{i=1}^{n} u_i\right)},$$

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

E) Generating Poisson processes



E.1) Homogeneous Poisson process with rate λ

Method 1

- Generate i.i.d. 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 $N \sim \text{Poisson}(\lambda \mathcal{T})$, the number of spikes on the interval $[0, \mathcal{T}]$.
  (In Matlab, use 'poissrnd')
- Draw $T_1, \ldots, T_N \sim \text{Uniform}([0, \mathcal{T}])$
  (In Matlab, use 'rand')
  The $T_1, \ldots, T_N$ are the spike times.

Why does Method 2 work?

The intuition is that a spike should not be more likely to occur at one time compared to another time (think of Bernoulli process).

More formally, Method 2 is based on the following theorem:

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 $U_1, \ldots, U_N \sim \text{Uniform}([0, \mathcal{T}])$ i.i.d.

$T_2 > T_1$ ordered

not ordered

We won't prove this theorem here.

E.2) Inhomogeneous Poisson process with rate $\lambda(t)$

- Let $\lambda_{max} = \max_t \lambda(t)$.

  Generate a homogeneous Poisson process with rate λ_{max} using either method in E.1).

- For $n = 1, \ldots, N$

thinning

Draw $U \sim \text{Uniform}([0, 1])$

If $U > \dfrac{\lambda(T_n)}{\lambda_{max}}$, 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 will not prove here why this is true.