Problem Set 4:

Stochastic Point Processes

Week 7, SoSe 2022 — Benoit Scholtes
Computational Neuroscience — & Stefan Rotter

Problem 1: The Poisson Process — 12 points

Recent lectures have shown the utility of using point processes to model and understand processes in the brain, such as spike trains. The Poisson point process is a completely random point process in that each point is stochastically independent to all other points.

Problem 1.1: Counting vs interval statistics — 4 points

There are different ways in which one can analyse or view point processes. One such way is via counting statistics, where the number of events (points) within non-overlapping intervals of length t are counted over the process, i.e. the number of spikes in 10 ms time intervals along a spike train. The counts are denoted as $\{N(t), t \geq 0\}$ and the distribution of the counts gives the probabilities of the counts (for intervals of length t) for the process $\mathbb{P}[N(t) = k]$. For a Poisson process with stationary intensity λ, the expected number of points in an interval of length t is $\mathbb{E}[N(t)] = \lambda t$. What is the distribution of spike counts for an interval of length t for the stationary Poisson process?

One can also look at point processes by looking at the statistics of the elapsed times between adjacent events (i.e. inter-spike interval density). Derive this distribution for the Poisson process with $\mathbb{E}[N(t)] = \lambda t$.

Hint: To do so, take these steps:

1. *If a spike of a Poissonian neuron occurs at $t_0 = 0$, and the next at some unknown t_1, what is then the relation between the probability $\mathbb{P}[N(t) = 0]$ from the counting statistics that there is no spike in interval $A = (0, t]$ for some random t, and the probability $\mathbb{P}[t_1 > t]$ from the interval statistics that elapsed time to the next spike $t_1 - t_0 = t_1$ is longer than t?*
2. *What is $\mathbb{P}[N(t) = 0]$ from the distribution you just found above.*
3. *What is the relation between $\mathbb{P}[t_1 > t]$ and the inter-spike interval density $f(t)$?*
4. *Substitute both of these into the relation you found in step one and solve for $f(t)$ (use the fundamental theorem of calculus!).*

Solution.

The distribution of spike counts for an interval of length t for the stationary Poisson process is exactly how the Poisson distribution

$$\mathbb{P}[N(t) = k] = \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

is defined, where here the intensity parameter (rate) is $\Lambda = \lambda t$. In other words, a Poisson point process has the property that its counting statistics follows a Poisson distribution.

Using the hints:

1. These must be equal, they are the same thing! $\implies \mathbb{P}[N(t)=0] = \mathbb{P}[t_1 > t]$
2. $\mathbb{P}[N(t)=0] = \frac{(\lambda t)^0}{0!} e^{-\lambda t} = e^{-\lambda t}$
3. If $f(t)$ is inter-spike interval density, then the probability $\mathbb{P}[t_1 > t]$ of the interval being longer than t is one minus the probability of the interval being less than or equal to t. This is expressed as $\mathbb{P}[t_1 > t] = 1 - F(t)$, where $F(t) = \mathbb{P}[t_1 \leq t] = \int_0^t f(s)\mathrm{d}s$ is the cumulative distribution function of the interval densities.
4. Thus:
$$\mathbb{P}[N(t)=0] = e^{-\lambda t} \quad \stackrel{!}{=} \quad \mathbb{P}[t_1 > t] = 1 - \int_0^t f(s)\mathrm{d}s$$

   Now, from the Fundamental Theorem of Calculus, $f(t) = \frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathrm{const.}}^t f(s)\mathrm{d}s$. Rearranging the previous line for $\int_0^t f(s)\mathrm{d}s$ and substituting this into the FTC:
$$\begin{aligned} f(t) &= \frac{\mathrm{d}}{\mathrm{d}t}\int_0^t f(s)\mathrm{d}s \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\left(1 - e^{-\lambda t}\right) \\ &= \lambda e^{-\lambda t} \end{aligned}$$

   Thus, the inter-spike interval density is a decaying exponential $f(t) = \lambda e^{-\lambda t}$.

Remember that in the second problem set, we modelled a channel's transitions as a discrete Markov processes with stationary opening and closing transition rates α and β, with individual transitions independent of all others. Note that the *collection* of opening events (or closing events) in time do **NOT** define a Poisson process. This is because adjacent opening events are separated by two exponentially distributed time intervals, first an channel-open interval followed by a channel-closed interval. The addition of these exponentially distributed variables gives a hypoexponentially distributed variable, such that the full interval between channel opening events is not exponential distributed as is for a Poisson process.

The open and closed dwell times distributions you found in the second problem set were geometric distributions, the discrete analogue to the exponential distribution. This is because the *individual* closing and opening events (i.e. looking at a channel opening event while the channel is **already** closed, so not including the previous channel-open interval) define two separate Poisson point processes (one each for opening and closing events). This is because when the channel is closed, for example, the probability of it opening at any moment is independent of time, given by the Markov transition rate above. This is what it means when we say that a Poisson process is memoryless. Exponential and geometric distributions are the only two interval density distributions which are memoryless.

Problem 1.2: Doubly stochastic process — 4 points

The Poisson point process considered above was stationary in that the rate λ did not change in time. Neurons however have firing rates which change in time $\lambda(t)$, depending on their input, etc. If the rate parameter $\lambda(t)$ is itself taken to be a random variable, then the resulting spike train is a *doubly stochastic* Poisson point process, in that the probability of counting some number of spikes $N(t)$ is dependent/*conditional* on the rate $\lambda(t)$ at that time: $\mathbb{P}[N(t) = k \mid \lambda(t)]$. Here, the rate $\lambda(t)$ of the process changes *stochastically* in continuous time, different to the nonstationary Poisson process where $\lambda(t)$ changes *deterministically* in time.

In the stationary Poisson process, we saw that the expected number of spikes observed in an interval of length t is $\mathbb{E}[N(t)] = \lambda t$. In the lecture, you saw that for the nonstationary Poisson process, the number of spikes N expected during an interval of length $(0, t]$ is just the integral of the rate during the interval $\mathbb{E}[N(0 \to t) \mid \lambda(t)] = \int_0^t \lambda(s)\mathrm{d}s \equiv \Lambda$. What is then the expected number of spikes during an interval $(0, t]$ in the doubly stochastic Poisson process? What about the variance of the doubly stochastic process?

Hint: What is the conditional *spike count distribution for the doubly stochastic process (*conditional *on the summed rate Λ during the count interval)?*

Also, use the laws of total expectation and total variance. These define the unconditional expectation $\mathbb{E}[X]$ and variance $\mathbb{V}ar[X]$ of a random variable X in terms of its conditional expectation $\mathbb{E}[X|Y]$ and variance $\mathbb{V}ar[X \mid Y]$. The laws are:

$$\mathbb{E}[X] = \mathbb{E}\big[\mathbb{E}[X \mid Y]\big], \quad \textit{and}$$
$$\mathbb{V}ar[X] = \mathbb{E}\big[\mathbb{V}ar[X \mid Y]\big] + \mathbb{V}ar\big[\mathbb{E}[X \mid Y]\big].$$

The proofs of the laws can be found here[1] and here[2], respectively, and may help to understand them.

Solution.

From lecture 5 on the nonstationary Poisson process, the number of spikes Λ expected $\mathbb{E}[N(t) \mid \lambda(t)]$ during an interval of length $(0, t]$ is just the integral of the *deterministic* rate during the interval:

$$\Lambda = \int_0^t \lambda(s)\mathrm{d}s$$

This is exactly the same for the doubly stochastic process, but as $\lambda(t)$ is itself stochastic and changes from trial to trial, Λ also becomes random. This then means that the spike count distribution, still being Poisson distributed, is *conditional* on this random parameter Λ:

$$\mathbb{P}[N(t) = k \mid \Lambda] = \frac{(\Lambda)^k}{k!}e^{-\Lambda},$$

such that the expected count *conditional* on Λ is indeed $\mathbb{E}[N(t) \mid \Lambda] = \Lambda$.

Expectation: Now we want to find the expected spike count *unconditional* on Λ (i.e.

[1] `https://en.wikipedia.org/wiki/Law_of_total_expectation`
[2] `https://en.wikipedia.org/wiki/Law_of_total_variance`

leaving it stochastic). From the law of total expectation $\mathbb{E}[N(t)] = \mathbb{E}\big[\mathbb{E}[N(t)\,|\,\Lambda]\big]$, the expected count during the doubly process is thus $\mathbb{E}[N(t)] = \mathbb{E}[\Lambda]$. This means that the expected spike count of the doubly stochastic Poisson process is the mean of the stochastic rate Λ.

Variance:

Plugging $\mathbb{E}[N(t)\,|\,\Lambda] = \Lambda$ and $\mathbb{V}\text{ar}[N(t)\,|\,\Lambda] = \Lambda$ for a (stationary) Poisson process into the law of total variance $\mathbb{V}\text{ar}[N(t)] = \mathbb{E}\big[\mathbb{V}\text{ar}[N(t)\,|\,\Lambda]\big] + \mathbb{V}\text{ar}\big[\mathbb{E}[N(t)\,|\,\Lambda]\big]$, we obtain:

$$\mathbb{V}\text{ar}[N(t)] = \mathbb{E}[\Lambda] + \mathbb{V}\text{ar}[\Lambda].$$

This means that the variance of a doubly stochastic Poisson process is increased from that of a stationary Poisson process $\mathbb{E}[\Lambda]$ (in the case that the means coincide) by the variance of the rate parameter $\mathbb{V}\text{ar}[\Lambda]$.

Problem 2: Renewal Processes — 4 points

You saw in Lecture 5 that neuron spike trains might not be well modelled by a stationary Poisson point process, motivating the need for nonstationary Poisson processes and expanding to renewal processes in general. The key objects of interest in renewal theory are the interval density (or holding times) $f(t)$ used already above, the survivor function $\mathcal{F}(t) = \mathbb{P}[T > t]$ giving the probability that the next spike occurs at some time T which is later than some specified time t, and the hazard function $h(t)$ defined as the likelihood that a neuron which hasn't already spiked for a time t will spike at that time.

Problem 2.1: Interval density vs. hazard function — 4 points

One problem of using the stationary Poisson process for spike trains is that it allows and frequently produces spikes very close in time, in contention with the observed refractory period of neurons. One way to rectify this is by imposing a dead time after each spike during which no spikes can occur. This immediately violates the independence of points, such that the process is longer Poissonian. We can, however, impose that the process is "Poissonian" *after* the dead time, in that there is then vanishing dependence on the previous spike. What we obtain for the interval density distribution for a neuron with dead time D is

$$f(t) = \begin{cases} 0 & \text{if } t \leq D, \\ \frac{\lambda}{a} e^{-\lambda t} & \text{if } t > D, \end{cases}$$

where λ is not the firing rate of the neuron but rather the rate *exclusive* of the dead times, and a is required to normalise the pdf such that $\int_0^\infty f(t)\,\mathrm{d}t = 1$. What is the expression for a?

Calculate the survivor function $\mathcal{F}(t)$ and the hazard function $h(t)$ of this renewal process.

Solution.

Firstly, for a we know that $\int_0^\infty f(t)\ \mathrm{d}t = 1$, thus:

$$\begin{aligned}
1 &= \int_0^\infty f(t)\ \mathrm{d}t \\
&= \int_D^\infty \frac{\lambda}{a} e^{-\lambda t}\ \mathrm{d}t
\end{aligned}$$

such that $a = \int_D^\infty \lambda e^{-\lambda t}\ \mathrm{d}t$. We can also calculate this more explicitly if we want:

$$\begin{aligned}
a &= \int_D^\infty \lambda e^{-\lambda t}\ \mathrm{d}t \\
&= -\left[e^{-\lambda t}\right]_D^\infty \\
&= e^{-\lambda D}
\end{aligned}$$

This means, the longer the refractory period, the higher than amplitude of the exponential.

For the survivor and hazard functions, we split the interval density function into its two pieces. For $f(t) = 0$ during $t \leq D$, we have

$$\begin{aligned}
\mathcal{F}(t) &= 1 - \int_0^t f(s)\ \mathrm{d}s \\
&= 1 - \int_0^t 0\ \mathrm{d}s \\
&= 1
\end{aligned}$$

for the survivor function, and

$$\begin{aligned}
h(t) &= \frac{f(t)}{\mathcal{F}(t)} \\
&= 0
\end{aligned}$$

for the hazard function. For $f(t) = \frac{\lambda}{a} e^{-\lambda t}$ during $t > D$, we have

$$\begin{aligned}
\mathcal{F}(t) &= 1 - \int_0^t f(s)\ \mathrm{d}s \\
&= \int_t^\infty f(s)\ \mathrm{d}s \\
&= [-\frac{\lambda}{a\lambda} e^{-\lambda s}]_t^\infty \\
&= \frac{1}{a} e^{-\lambda t}
\end{aligned}$$

for the survivor function, and

$$\begin{aligned}h(t) &= \frac{f(t)}{\mathcal{F}(t)} \\ &= \frac{\lambda/a\ e^{-\lambda t}}{1/a\ e^{-\lambda t}} \\ &= \lambda\end{aligned}$$

for the hazard function. Thus, the full survivor function for the process is

$$\mathcal{F}(t) = \begin{cases} 1 & \text{for } t \leq D, \\ e^{-\lambda t} & \text{for } t > D, \end{cases}$$

while the complete hazard function is

$$\mathcal{F}(t) = \begin{cases} 0 & \text{for } t \leq D, \\ \lambda & \text{for } t > D. \end{cases}$$

Problem 2.2: The waiting time paradox **4 points**

Let's see how stochastic point processes can be applied to the macroscopic world. Assume you having a working bus system where a bus comes every 15 minutes at a certain bus stop without fail. You've had complaints from commuters, however, that the average waiting time for buses is too long and that they'd like it shortened. The problem is that you don't have the budget to buy more buses. The best idea in your mind would be to allow buses to make their rounds as quickly as possible, no longer sticking to the strict 15 minute timetable. Due to your naivety and desire for simplification, you assume that these buses arriving at any given bus stop represent a Poisson process with rate of $\lambda = 1/15\ \text{min}^{-1}$ (since you still have the same number of buses). Will the mean waiting time be reduced under this Poissonian system?

Hint: You should find the following integral useful

$$\int_0^\infty xe^{-ax}\,\mathrm{d}x = \frac{1}{a^2}.$$

Solution.

First we need to calculate the mean waiting time for buses arriving on the 15 minute schedule. Here, we have to assume that the time that people arrive t_{arrival} at the bus stop around this schedule is random and with uniform probability, i.e. a person has an equal probability of arriving 1 or 8 or any $x < 15$ minutes before a bus is scheduled to arrive. This means that the waiting times $w = t_{\text{bus}} - t_{\text{arrival}}$ are themselves uniformly distributed on the interval $(0, 15)$. The uniform distribution $U(a, b)$ has pdf $f(x) = 1/(b-a)$ for $x \in (a, b)$ and 0 elsewhere. We can thus calculate the mean, being the expected values

$E[w]$:

$$\begin{aligned}
E[w] &= \int_0^\infty s\ f(s)\ \mathrm{d}s \\
&= \int_0^{15} s\ \frac{1}{15-0}\ \mathrm{d}s \\
&= \frac{1}{15} \cdot \frac{1}{2}[s^2]_0^{15} \\
&= 7.5,
\end{aligned}$$

such that the mean waiting time of the current scheduled system is half that of the time between bus arrivals.

Now, since the Poisson process is memoryless (in that a bus has the same probability of arriving at any point in time, irrespective of when the last buses came OR when the commuter arrived to the bus stop/how commuter arrivals are distributed), then the arrival of buses still follows the same Poisson process for the commuter. Thus, the waiting times follow the interval density $f(x) = \lambda e^{-\lambda x}$ with parameter $\lambda = 1/15$, such that the mean waiting time is

$$\begin{aligned}
E[w] &= \int_0^\infty s\ \lambda e^{-\lambda s}\ \mathrm{d}s \\
&= \lambda(\frac{1}{\lambda^2}) \\
&= \frac{1}{\lambda} \\
&= 15,
\end{aligned}$$

where the provided integral was used for the second step. Thus, we find that while buses come at the same rate for the scheduled and Poissonian systems, and thus have the same mean number of buses arriving per hour (or whatever time interval), the mean waiting time for the Poissonian system is double that of the strict schedule!