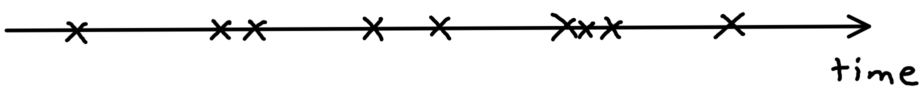
Poisson Process (with probability review)

In this note, we consider an important random process called Poisson process. This process is a popular model for customer arrivals or calls requested to telephone systems.

We start by modeling Poisson Process as a random arrangement of "marks" (denoted by "x") on the time line. These marks may indicate the time that customers arrive or the time that call requests are made.



We will focus on one kind of Poisson process:

homogeneous Poisson process

From now on, when we say "Poisson process", what we mean is "homogeneous Poisson process".

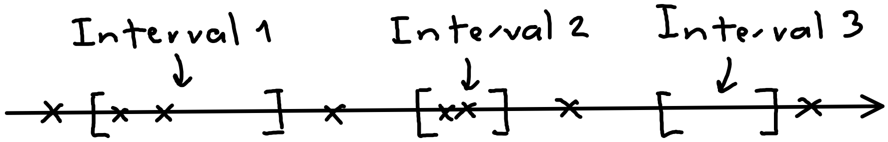
The first property of Poisson process that you should remember is that
there is only <u>one</u> parameter for Poisson process. This parameter is the rate or intensity of arrivals (the average number of arrivals per unit time).
We use λ to denote this parameter.

If λ is a constant, the Poisson process is homogeneous.

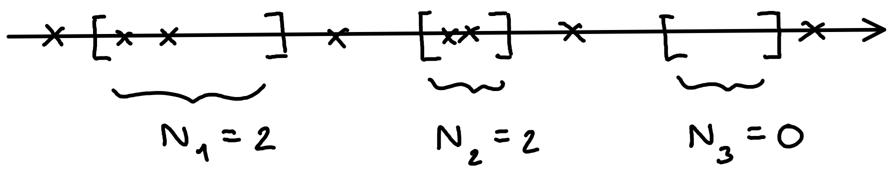
If λ is a function of time, say $\lambda(t)$, the Poisson process is non-homogeneous.

Our λ is constant because we focus on homogeneous Poisson process.

So, how can this λ control the Poisson Process? The key idea is that the Poisson process (PP) is as random/unstructured as a process can be. Therefore, if we consider many non-overlapping intervals on the time-line shown below,



and count the number of arrivals in these intervals.



Then, the numbers N_1, N_2, N_3 should be independent; that is knowing the value of N_1 does not tell us anything at all about what N_2 and N_3 will be. This is what we are going to take as a vague definition of the "complete randomness" of the Poisson process.

So now, we have one more property of PP:

The number of arrivals in non-overlapping intervals are independent.

By saying something are independent, of course, we mean it in terms of probability. Note that the numbers N_1, N_2, N_3 above are <u>random</u>. Because they are counting the number of arrivals, we know that they can be any non-negative integers:

$$0, 1, 2, 3, \ldots\ldots$$

Because we don't know their exact values, we

Because we don't know their exact values, we describe them via the likelihood or <u>probability</u> that they will take one of these values.
For example, for N_1, we describe it by

$$P[N_1=0],\ P[N_1=1],\ P[N_1=2],\ \dots$$

where
$P[N_1=k]$ is the probability that N_1 takes the value k.

The above notation is a bit long. So, we define

$$p_{N_1}(k) = P[N_1=k].$$

This $p_{N_1}(\cdot)$ is then a function of k which tells the probability that N_1 will take a particular value k. We call p_{N_1} the <u>probability mass function</u> (pmf) of N_1

At this point, we don't know much about $p_{N_1}(k)$ except that they are between 0 and 1 and the sum

$$\sum_{k=0}^{\infty} p_{N_1}(k) = 1.$$

These two are the necessary and sufficient properties of any pmf.

When we say that N_1 and N_2 are independent, it means that

$$P[N_1=k \text{ and } N_2=m]$$

(which is the probability that $N_1=k$ and $N_2=m$) can be written as the product

$$p_{N_1}(k) \times p_{N_2}(m).$$

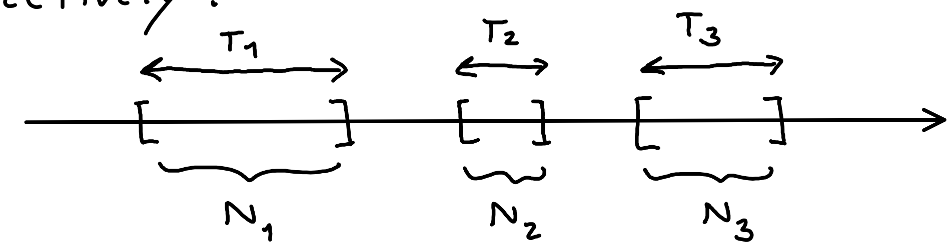
Do we know anything else about N_1, N_2, N_3?

What about the λ? Can we connect λ to N_1, N_2, N_3?

Recall that λ is the average number of arrival per unit time.

So, if $\lambda = 5$ arrivals/hour then we expect that N_1, N_2, N_3 should statistically agree with this λ. How?

Let's first be more specific about the time duration of the intervals that we have earlier. Suppose their lengths are T_1, T_2, T_3 respectively:



Then, you should expect that

$$\mathbb{E}N_1 = \lambda T_1$$

$$\mathbb{E}N_2 = \lambda T_2$$

$$\mathbb{E}N_3 = \lambda T_3$$

Recall that $\mathbb{E}N_1$ is the (average) expectation of the random variable N_1.

$$\mathbb{E}N_1 = \sum_k k \times P[N_1 = k].$$

"value" weighted by "frequency of occurrence"

For example, suppose $\lambda = 5$ arrivals/hour and $T_1 = 2$ hour.

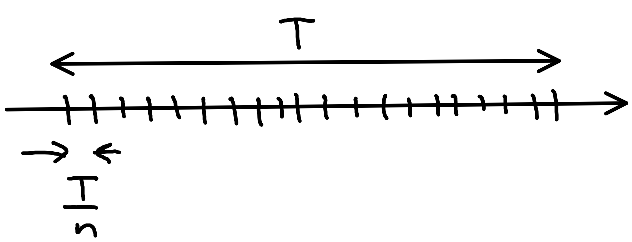
Then you would get about $\lambda \times T_1 = 10$ arrivals.

during the first interval. Of coure, the number of arrivals is random. So, this number 10 is an average or the expected number, not the actual value.

In conclusion, we now know more about PP. For any interval of length T, the expected nunber of arrivals in this interval is given by

$$\mathbb{E}N = \lambda T.$$

The next key idea is to consider a small interval. Imagine dividing a time interval of length T into n equal slots.



Then each slot would be a time interval of duration $\frac{T}{n}$.

If $T = 20$ hr. and $n = 10{,}000$, then each small interval would have length

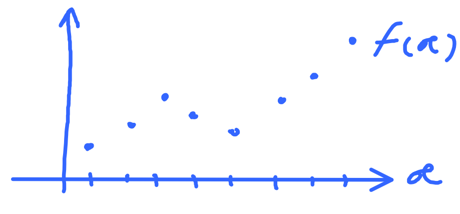
$$\frac{T}{n} = \frac{20}{10{,}000} = 0.002 \text{ hour.}$$

Why do we consider small interval?
The key is that as the interval becomes very small, then it is extremely unlikely that there will be more than 1 arrivals during this small amout of time!

The above statement becomes more accurate as we increase the value n which decreses the length of each small interval even further!

What we are doing here is to approximate a continuous-time process by a discrete-time process.

(You also do this when you plot a graph of any function $f(x)$. You evaluate the values of the function for many values of x where the x are close enough such that nothing surprising can happen in between.



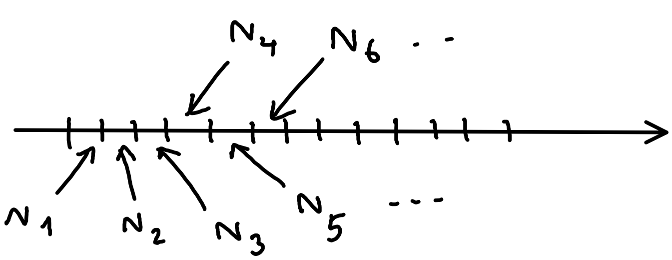
)

If we want to be more vigorous, we would have to bound the error from such approximation, and show that the error disappear as $n \to \infty$. I will not do that here.

What do we gain by considering discrete-time approximation by small intervals?

When the interval is small enough, we can assume that at most 1 arrival occurs.

So now let N_1 be ~~#~~ arrivals in slot 1
N_2 be " " 2
N_3 be " " 3
and so on as shown below:



Then, the N_i are all Bernoulli random variable!
(This is because they can only be 0 and 1.)

In which case, the pmf for the N_i reduces to only two values to only two values

$$P[N_i = 0] \text{ and } P[N_i = 1]$$

We also know that the average number of arrivals should be

$$\mathbb{E} N_i = \lambda \times \underbrace{\frac{T}{n}}_{\text{length of each slot}}$$

So, all N_i are <u>Bernoulli</u> r.v. with the same average!! Also, they are all independent because they come from non-overlapping intervals.

For Bernoulli r.v. X, the average is

$$\mathbb{E} X = 0 \times P[X=0] + 1 \times P[X=1]$$
$$= P[X=1]$$

We use $p_0 = P[X=0]$ and $p_1 = P[X=1]$

to simplify the notation for Bernoulli r.v.

Hence,

$$\mathbb{E} X = p_1$$

Of course, we should also know that $p_0 + p_1 = 1$. In other words, if we know $\mathbb{E} X$ of a Bernoulli random variable, we also know its pmf:

$$p_1 = P[X=1] = \mathbb{E} X$$
$$p_0 = P[X=0] = 1 - \mathbb{E} X.$$

For our discrete-time approx. of PP, we now know that all of the $N_1, N_2, \ldots$ have the same expectation

$$\mathbb{E} N_1 = \mathbb{E} N_2 = \mathbb{E} N_3 = \cdots = \lambda \frac{T}{n}$$

For $\lambda = 5$, $T = 20$, $n = 10{,}000$

$$\lambda \frac{T}{n} = 0.01 \text{ arrival}.$$

So, the pmf's of all $N_1, N_2, \ldots$ are all the same!!

They are governed by

$$p_1 = \lambda \frac{T}{n} \quad \text{and} \quad p_0 = 1 - \lambda \frac{T}{n}$$

The probability that there will be an arrival in the small interval

We say that $N_1, N_2, N_3, \ldots$ are i.i.d.

(independent and identically distributed)

At this point, you can use MATLAB to generate a PP using this discrete time approx.

First, we fix the end time T of the simulation.

(ex. T = 20 hr.)

Then, we divide T into n slots

(ex. n = 10,000)

For each slot, only two cases can happen:

1 arrival

or

no arrival

So, generate Bernoulli r.v. for each slot with

$p_1 = \lambda \times \frac{T}{n}$ (if $\lambda = 5$ arrivals/hr, then $p_1 = 0.01$)

To do this for 10,000 slots at the same time, we can use

rand(1, n) < p_1

or

binornd$(1, p_1, 1, n)$