Poisson processes

Definition

A *Poisson process with rate* λ is a counting process $(N(s))_{s\in[0,\infty)}$ with $N(0) = 0$, whose inter-arrival times are i.i.d. exponential random variables with rate λ.

Poisson Process

Theorem

$(N(s))_{s\in[0,\infty)}$ is a Poisson process with rate λ if and only if it is a counting process such that

1. $N(0) = 0$,
2. *it has independent increments;*
3. $N(t+s) - N(s) \sim \mathit{Poisson}(\lambda t)$.

Poisson Process

Theorem

$(N(s))_{s\in[0,\infty)}$ is a Poisson process with rate λ if and only if it is a counting process such that:

1. $N(0) = 0$;
2. *it has independent increment;*
3. *it has stationary increments;*
4. $\lim_{h\to 0} \frac{P(N(h)=1)}{h} = \lambda$ *and* $\lim_{h\to 0} \frac{P(N(h)\geq 2)}{h} = 0$.

Thinning

An arrival process, $(N(s))_{s\in[0,\infty)}$ is Poisson with rate λ. Assume that the arrivals can be of k different types, specified by a sequence of iid random variables $\{Y_i\}_{i=1}^{\infty}$, taking values in $\{1, 2, 3, \dots, k\}$, with probability mass function $P(Y_i = j) = p_j$. Let these random variables be **independent of** $(N(s))_{s\in[0,\infty)}$. Let $N_j(t)$ be the arrivals before time t that are of type j:

$$N_j(t) = \sum_{i=1}^{N(t)} 1_{\{Y_i = j\}}.$$

Theorem

$\{(N_j(t))_{t\in[0,\infty)}\}_j$ are independent Poisson processes with respective rates λp_j.

Superposition

$t_1 < t_2 < t_3 < t_4$ $\quad N(t_2) - N(t_1) = \sum_{j=1}^{k} N_j(t_2) - N_j(t_1)$

$N(t_4) - N(t_3) = \sum_{j=1}^{k} N_j(t_4) - N_j(t_3)$

Theorem

Suppose that $N_1(t), \ldots, N_k(t)$ are independent Poisson processes with rates $\lambda_1, \ldots, \lambda_k$, respectively. Then

$$N(t) = N_1(t) + \cdots + N_k(t),$$

is a Poisson process with rate $\lambda_1 + \cdots + \lambda_k$.

Proof.

We simply check the second definition of Poisson process.

1. $N(0) = N_1(0) + \cdots + N_k(0) = 0.$ ✓
2. Independent increments follows from that of the other processes
3. Increments are Poisson since the sum of Poisson independent r.v. is Poisson

$$N(t+s) - N(t) = \sum_{j=1}^{k} \overbrace{N_j(t+s) - N_j(t)}^{\sim \mathrm{Pois}(\lambda_j s) \ \text{IND.}} \sim \mathrm{Pois}\Big(\sum_{j=1}^{k} \lambda_j s\Big)$$

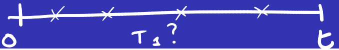
□

Example

Assume people from city A and people from city B arrive at a stadium according to two independent Poisson processes $(N_A(t))_{t=0}^{\infty}$ and $(N_B(t))_{t=0}^{\infty}$ with rates 10 and 20 people per hour, respectively. The ticket seller only cares about how many tickets are sold, he does not care about the buyer comes from city A or B. That is, the ticket seller only cares about the process $N(t) = N_A(t) + N_B(t)$, which corresponds to the total number of tickets sold by time t. By superposition, $(N(t))_{t\in[0,\infty)}$ is a Poisson process with rate 10+20=30 per hour. Hence, for example we have

$$P(\underbrace{N(2)}_{\text{Pois}(30\cdot 2)} = 4) = e^{-2\cdot 30}\frac{(2\cdot 30)^4}{4!}$$

Conditioning



Assume that we have had only one arrival up to time t. What is the distribution of the arrival time T_1, given this information? It is uniform on $[0, t]$.

$s < t$ DEF.

$$P(T_1 > s | N(t) = 1) \overset{\downarrow}{=} \frac{P(N(s) = 0, N(t) = 1)}{P(N(t) = 1)} = \frac{P(N(s) = 0, N(t) - N(s) = 1)}{P(N(t) = 1)}$$

independent increments $\rightarrow$
$$= \frac{P(N(s) = 0)P(N(t) - N(s) = 1)}{P(N(t) = 1)}$$

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

$$= 1 - \frac{s}{t} = P(U > s),$$

$P(U > s) = \frac{t-s}{t}$

where U is a uniform random variable on $[0, t]$. Hence, T_1 is uniformly distributed on $[0, t]$ given that $N(t) = 1$! No matter λ.

Conditioning $N(t) = n$ $\sim$ $U_1, U_2, \dots U_n$ i.i.d.

Something more general holds: $(T_1, T_2, T_3, \dots T_n)|_{N(t)=n}$ $\cancel{\sim}$ $(U_1, U_2, \dots, U_n)$

ORDINATI i.i.d. UNIF.

Theorem

*Given that $N(t) = n$, the set $\{T_1, T_2, \dots, T_n\}$ of arrival times in $[0, t]$ is distributed as a set $\{U_1, U_2, \dots, U_n\}$ of n **independent uniform** random variables on $[0, t]$.*

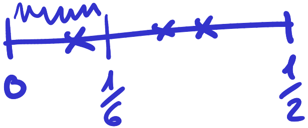
Corollary

If $s < t$ and $n \geq 0$, then given that $N(t) = n$ the random variable $N(s)$ is distributed as a binomial random variable $Bin(n, s/t)$.

Proof.

The result follows from the fact that, given that $N(t) = n$, the set of arrival times up to t is distributed as a set of n independent random variables, each one with probability s/t of being smaller than s. $N(t) = n$

Example



Assume customers arrive to a shop with rate of 5 customers per hour. Given that exactly 3 customers arrive in the first 30 minutes, what is the probability that exactly 2 customers entered the shop in the first ten minutes? In symbols, what is $P(N(1/6)=2|N(1/2)=3)$?

$$P(N(1/6)=2|N(1/2)=3)=\binom{3}{2}\left(\frac{1/6}{1/2}\right)^2\left(1-\frac{1/6}{1/2}\right)=3\left(\frac{1}{3}\right)^2\frac{2}{3}=\frac{2}{9}.$$

$$N\left(\tfrac{1}{6}\right) \mid N\left(\tfrac{1}{2}\right)=3 \sim \mathrm{Bin}\left(3, \frac{1/6}{1/2}\right)$$

INCR. IND.

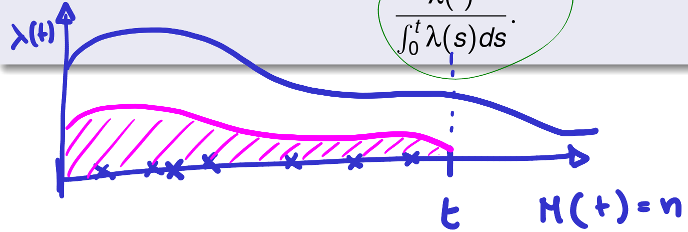
$$P\left(N\left(\tfrac{1}{2}\right)=3 \mid N\left(\tfrac{1}{6}\right)=2\right) = P\left(N\left(\tfrac{1}{2}\right)-N\left(\tfrac{1}{6}\right)=1 \mid \cancel{N\left(\tfrac{1}{6}\right)=2}\right)$$

Conditioning and non-homogeneous Poisson process

Theorem

Consider a non-homogeneous Poisson process $\{M(t) : t \geq 0\}$ with rate $\lambda(\cdot)$. Given that $M(t) = n$, the set $\{T_1, T_2, \ldots, T_n\}$ of arrival times in $[0, t]$ is distributed as a set $\{U_1, U_2, \ldots, U_n\}$ of n ***independent*** *random variables on $[0, t]$ with density*

$$\frac{\lambda(\cdot)}{\int_0^t \lambda(s)ds}.$$



3° SIMULATION STRATEGY

1. $M(t) \sim \mathrm{Pois}\left(\int_0^t \lambda(s)ds\right)$
2. SIMULARE INSIEME DI ARRIVI

Extensions

We do not have time to look at more extensions of the Poisson process, but I list here some of the most famous ones (no need to rember them at the exam!)

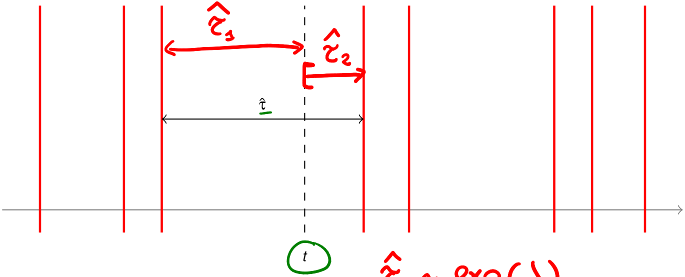
- Spatial Poisson process (the arrivals are connected with a random location in space - bombing during WW2, earthquake locations, lightnings, forest growing, etc.);
- Compound Poisson process (the arrival size is a random variable - arrivals to restaurants, museums, insurance, etc.);
- Poisson point process (more theoretical, it is a way to distrubute random points on a general topological space)

The inspection paradox

The lenght of a pinned interarrival interval

$t_1, t_2, \ldots$ INTERARRIVAL TIMES $\sim exp(\lambda)$

Consider a Poisson process with rate λ. Fix a time t different from 0, and let $\hat{\tau}$ be the length of the interarrival time containing t.



$\hat{\tau}_2 \sim exp(\lambda)$

What is $\mathbb{E}[\hat{\tau}]$?

$E[\hat{\tau}] \neq \frac{1}{\lambda}$

$E[\hat{\tau}] = E[\hat{\tau}_1] + E[\hat{\tau}_2] = E[\hat{\tau}_1] + \frac{1}{\lambda} > \frac{1}{\lambda}$

The inspection paradox

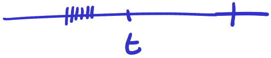
In general, whenever we have a renewal process with interarrival intervals of length t_i, if we fix a large enough t and define by $\hat{\tau}$ the length of the interarrival time containing t, then

$$\mathbb{E}[\hat{\tau}] > \mathbb{E}[t_i].$$

$E[\hat{\tau}] = E[t_i]$ NEL CASO IN CUI TUTTI I t_i SIANO DETERMINISTICI

Example

Let's make a trivial example: let $(t_i)_{i=1}^{\infty}$ be i.i.d. with



$$t_i = \begin{cases} 1/9 & \text{with probability} 9/10 \\ 10 & \text{with probability} 1/10. \end{cases}$$

Then,

$$\mathbb{E}[t_i] = \frac{1}{9} \cdot \frac{9}{10} + 10 \cdot \frac{1}{10} = \frac{11}{10}.$$

However, on average the short intervals cover $1/11$ of the whole time (you can use the theory of renewal reward processes to prove it), so we expect our fixed time t is in a long interval with much more probability! As a consequence, if t is large, we expect

$$\mathbb{E}[\hat{\tau}] \approx \frac{1}{9} \cdot \frac{1}{11} + 10 \cdot \frac{10}{11} > \mathbb{E}[t_i].$$

The paradox is everywhere!

Assume that there are two classes, one with 30 students and another one with 10 students. then, the average size of the classes is 20.
However if you interview a randomly chosen student, the *perceived* average size is larger: if A is the student's answer then

$$\mathbb{E}[A] = \boxed{30} \cdot \frac{30}{40} + \boxed{10} \cdot \frac{10}{40} = 25 > 20.$$

interviewing a student is similar to fix a time point t and ask it: how large is the interarrival time containing you? Picking a time point pertaining to a longer interarrival interval is more likely!

When transport companies complain that the trains or airplanes or buses are mostly empty, and passengers complain that they are often uncomfortably crowded, they may be both right! Simply, more people experience the crowd.