Recap: Process with stationary and independent increments

- A (continuous-time) random process $\{N(t),\ t \geq 0\}$ has independent increments if for all $t_1 < t_2 \leq t_3 < t_4$, $N(t_2) - N(t_1)$ is independent of $N(t_4) - N(t_3)$
- In particular, for all $s > t > 0$, $N(s) - N(t)$ is independent of $N(t) - N(0)$
- The process has stationary increments if $N(t_2) - N(t_1)$ has the same distribution as $N(t_2 + \tau) - N(t_1 + \tau)$, $\forall \tau, \forall t_2 > t_1$
- A random process $\{N(t),\ t \geq 0\}$ is called a counting process if $N(t)$ is non-negative integer-valued and is non-decreasing

Recap: Poisson Process

- We can define it in two ways
- **Definition 1** $\{N(t), t \geq 0\}$ is a counting process with
  1. $N(0) = 0$
  2. The process has stationary and independent increments
  3. $Pr[N(t) = n] = e^{-\lambda t} \, \frac{(\lambda t)^n}{n!}, \; n = 0, 1, \cdots$
- **Definition 2** $\{N(t), t \geq 0\}$ is a counting process with
  1. $N(0) = 0$
  2. The process has stationary and independent increments
  3. $Pr[N(h) = 1] = \lambda h + o(h)$ and
     $Pr[N(h) \geq 2] = o(h)$
- The two definitions are equivalent.

Recap: n^{th} order distributions

- The first order distribution of the process is: $N(t) \sim \text{Poisson}(\lambda t)$
- This, along with stationary and independent increments property determines all distributions

$$Pr[N(t_1) = n_1, N(t_2) = n_2, N(t_3) = n_3]$$
$$= Pr[N(t_1) = n_1]\; Pr[N(t_2) - N(t_1) = n_2 - n_1]$$
$$Pr[N(t_3) - N(t_2) = n_3 - n_2]$$
$$= Pr[N(t_1) = n_1]\; Pr[N(t_2 - t_1) = n_2 - n_1]\; Pr[N(t_3 - t_2) = n_3 - n_2$$

where we assumed $t_1 < t_2 < t_3$

Recap: mean and autocorrelation

$$
\begin{aligned}
\eta_N(t) &= E[N(t)] = \lambda\, t \\
R_N(t_1, t_2) &= \lambda^2 t_1 t_2 + \lambda\, \min(t_1, t_2) \\
&\Rightarrow \quad \text{not stationary}
\end{aligned}
$$

Recap: Inter-arrival or waiting times

- Let T_1 denote the time of first event and let T_n denote the time between n^{th} and $(n-1)st$ events.

$$Pr[T_1 > t] = Pr[N(t) = 0] = e^{-\lambda t}$$

- T_n are iid exponential with parameter λ
- The time of n^{th} event is

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

- Since T_i are iid, exponential, S_n is Gamma with parameters n, λ

Recap: Conditional distribution of event times

- Let $0 < s < t$.

$$Pr[S_1 \leq s | N(t) = 1] \;\; = \;\; \frac{s}{t}$$

- Conditioned on $N(t) = 1$, S_1 is uniform over $[0, t]$

- Let $S_1, \cdots, S_n$ be the times of the first n events.
- We want the conditional joint density of $S_1, \cdots, S_n$ conditioned on $N(t) = n$.
- Note that the S_i have to satisfy $S_1 < S_2 < \cdots < S_n$.
- We showed that the conditional joint density of $S_1, \cdots, S_n$ conditioned on $N(t) = n$, would be same as the order statistics of n iid random variables uniform over $[0,\ t]$.

- Take $t_i, 1 \leq i \leq n$ satisfying $0 < t_1 < t_2 < \cdots < t_n < t$.
- Let h_i be small positive numbers such that $t_i + h_i < t_{i+1}, \forall i$.

$$
\begin{aligned}
& Pr[t_i \leq S_i \leq t_i + h_i, i = 1, \cdots, n \mid N(t) = n] \\
&= \frac{Pr[t_i \leq S_i \leq t_i + h_i, i = 1, \cdots, n, N(t) = n]}{Pr[N(t) = n]} \\
&= \frac{Pr[1 \text{ event in each } [t_i,\ t_i + h_i], 1 \leq i \leq n,\ 0 \text{ in rest of } [0,\ t]]}{Pr[N(t) = n]} \\
&= \frac{\left(\prod_{i=1}^{n} \lambda h_i e^{-\lambda h_i}\right)\ e^{-\lambda(t - h_1 - \cdots - h_n)}}{((\lambda t)^n / n!) e^{-\lambda t}} \\
&= \frac{n!\ h_1 \cdots h_n}{t^n}
\end{aligned}
$$

- Thus we have for $0 < t_1 < \cdots < t_n < t$,

$$\frac{Pr[t_i \leq S_i \leq t_i + h_i, i = 1, \cdots, n \mid N(t) = n]}{h_1 \cdots h_n} = \frac{n!}{t^n}$$

- If we now take limit as all h_i go to zero, the LHS above would be the conditional joint density of $S_1, \cdots, S_n$ conditioned on $N(t) = n$.
- Thus

$$f_{S_1 \cdots S_n | N(t)}(t_1, \cdots, t_n | n) = \frac{n!}{t^n}, \ 0 < t_1 < \cdots < t_n < t$$

- Let $X_1, \cdots, X_n$ be iid continuous random variables with common density f_x.
- Recall that $X_{(k)}$ denotes the k^{th} smallest of them.
- Then the joint density of $X_{(1)}, \cdots, X_{(n)}$ is given by

$$f(x_1, \cdots, x_n) = n! \prod_{i=1}^{n} f_x(x_i), \; x_1 < \cdots < x_n$$

- If X_i are uniform over $[0, \; t]$

$$f(t_1, \cdots, t_n) = \frac{n!}{t^n}, \; 0 < t_1 < \cdots < t_n < t$$

Examples

- We look at a few simple example problems using Poisson process.

$$\begin{aligned} E[N(4) - N(2) | N(1) = 3] &= E[N(4) - N(2)] \\ &= E[N(2) - 0] = 2\lambda \end{aligned}$$

- Another example;

$$E[S_4] = E\left[\sum_{i=1}^{4} T_i\right] = \frac{4}{\lambda}$$

- Vehicles on a road come as a Poisson process. A man would cross the road if he can see that no vehicle comes to that point for the next T time units. What is the probability that he would have zero waiting time?
  – same as the prob. of no arrivals in $[t - T,\ t]$.

- The memoryless property of exponential rv can be useful

$$Pr[S_3 > t | N(1) = 2] = \begin{cases} 1 & \text{if } \; t < 1 \\ e^{-\lambda(t-1)} & \text{if } \; t \geq 1 \end{cases}$$

- We can explicitly derive this.

- Taking $t > 1$, we have

$$
\begin{aligned}
Pr[S_3 > t | N(1) = 2] &= \frac{Pr[S_3 > t, \ N(1) = 2]}{Pr[N(1) = 2]} \\
&= \frac{Pr[2 \text{ event in } (0, \ 1], \ 0 \text{ in } (1, \ t)]}{Pr[N(1) = 2]} \\
&= \frac{Pr[2 \text{ event in } (0, \ 1]] \ Pr[0 \text{ in } (1, \ t)]}{Pr[2 \text{ event in } (0, \ 1]]} \\
&= e^{-\lambda(t-1)}
\end{aligned}
$$

- Here is another example

$$E[S_4 | N(1) = 2] = 1 + E[S_2] = 1 + \frac{2}{\lambda}$$

We calculate $Pr[S_4 > t | N(1) = 2]$ and use it to find the above expectation

- Taking $t > 1$,

$$\begin{aligned}
Pr[S_4 > t | N(1) = 2] &= \frac{Pr[S_4 > t,\ N(1) = 2]}{Pr[N(1) = 2]} \\
&= \frac{Pr[2 \text{ event in } (0,\ 1],\ 0 \text{ or } 1 \text{ in } (1,\ t)]}{Pr[N(1) = 2]} \\
&= \frac{Pr[2 \text{ event in } (0,\ 1]]\ Pr[0 \text{ or } 1 \text{ in } (1,\ t)]}{Pr[2 \text{ event in } (0,\ 1]]} \\
&= e^{-\lambda(t-1)} + \lambda(t-1)e^{-\lambda(t-1)}
\end{aligned}$$

- Recall that if X is a non-negative continuous rv, then

$$EX = \int_0^\infty (1 - F_X(x))\ dx$$

- Since S_4 is non-negative, we can use this

- We derived
$$Pr[S_4 > t | N(1) = 2] = e^{-\lambda(t-1)} + \lambda(t-1)e^{-\lambda(t-1)}, \ t > 1$$
What is its value for $0 < t < 1$?
- Using this

$$\begin{aligned}
E[S_4 | N(t) = 2] &= \int_0^1 1 \, dt + \int_1^\infty \left(e^{-\lambda(t-1)} + \lambda(t-1)e^{-\lambda(t-1)}\right) \, dt \\
&= 1 + \int_1^\infty e^{-\lambda(t-1)} \, dt + \int_1^\infty \lambda(t-1)e^{-\lambda(t-1)} \, dt \\
&= 1 + \int_0^\infty e^{-\lambda t} \, dt \ + \int_0^\infty \lambda \, t \, e^{-\lambda t} \, dt \\
&= 1 + \frac{1}{\lambda} + \frac{1}{\lambda} = 1 + \frac{2}{\lambda}
\end{aligned}$$

Example

- Given a specific T_0 we want to guess which is the last event before T_0.
- Consider a strategy: we will wait till $T_0 - \tau$ and pick the next event as the last one before T_0.
- The probability of winning for this is

$$Pr[\text{exactly 1 event in } (T_0 - \tau,\ T_0)] = \lambda\tau e^{-\lambda\tau}$$

- We pick τ to maximize this

$$\lambda e^{-\lambda\tau} - \lambda^2 \tau e^{-\lambda\tau} = 0 \quad \Rightarrow \quad \tau = \frac{1}{\lambda}$$

- Intuitively reasonable because expected inter-arrival time is $\frac{1}{\lambda}$

- Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ
- Suppose each event can be one of two types – Typ-I or Typ-II
  - $N_1(t) =$ number of Typ-I events till t
  - $N_2(t) =$ number of Typ-II events till t
  - Note that $N(t) = N_1(t) + N_2(t), \; \forall t$
- Suppose that, independently of everything else, an event is of Typ-I with probability p and Typ-II with probability $(1-p)$

  Theorem: $\{N_1(t), t \geq 0\}$ and $\{N_2(t), t \geq 0\}$ are Poisson processes with rate λp and $\lambda(1-p)$ respectively, and they are independent

$$
\begin{aligned}
& Pr[N_1(t) = n, N_2(t) = m] \\
& \quad = \sum_k Pr[N_1(t) = n, N_2(t) = m \mid N(t) = k] \; Pr[N(t) = k] \\
& \quad = Pr[N_1(t) = n, N_2(t) = m \mid N(t) = m + n] \; Pr[N(t) = m + n] \\
& \quad = \; {}^{m+n}C_n \; p^n \; (1-p)^m \; e^{-\lambda t} \; \frac{(\lambda t)^{m+n}}{(m+n)!} \\
& \quad = \frac{(m+n)!}{m! \; n!} \; p^n \; (1-p)^m \; e^{-\lambda(p+1-p)t} \; \frac{(\lambda t)^m (\lambda t)^n}{(m+n)!} \\
& \quad = \frac{(\lambda p t)^n}{n!} \; e^{-\lambda p t} \; \frac{(\lambda (1-p) t)^m}{m!} \; e^{-\lambda(1-p)t}
\end{aligned}
$$

- This shows that $N_1(t)$ and $N_2(t)$ are independent Poisson

- The interesting issue here is that $N_1(t)$ and $N_2(t)$ are independent.
- Suppose customers arrive at a bank as a Poisson process with rate 12 per hour.
- Independently of everything, an arriving customer is male or female with equal probability.
- We expect equal number of male and female customers; each have rate 6 per hour
- Q: Given that on some day 6 male customers came in the first half hour, what is the expected number of female customers in that half hour?
- The answer is 3 because the two processes are independent

- The theorem is easily generalized to multiple types for events
- Consider Poisson process with rate λ
- Suppose, independently of everything, an event is Type-i with probability p_i, $i = 1, \cdots K$.
- Note we have $\sum_{i=1}^{K} p_i = 1$
- Let $N_i(t)$ be the number of Type-i customers till t
- Then, these are independent Poisson processes with rates λp_i, $i = 1, \cdots, K$
- This is sometimes referred to as thinning of a Poisson process

- Superposition of independent Poisson processes also gives Poisson process.
- If N_1 and N_2 are independent Poisson processes with rates λ_1 and λ_2 then $N(t) = N_1(t) + N_2(t)$ is a Poisson process with rate $\lambda_1 + \lambda_2$
- We know that sum of independent Possion rv's is Poisson

Example

- Suppose number of radioactive particles emitted is Poisson with rate λ.
- We are counting particles using a sensor
- Suppose (independent of everything) an emitted particle is detected by our sensor with probability p
- Given that we detected K particles till t what is the expected number of particles emitted?
- Let these processes be $N(t), N_1(t), N_2(t)$

$$\begin{aligned} E[N(t)|N_1(t) = K] &= E[N_1(t) + N_2(t)|N_1(t) = K] \\ &= K + E[N_2(t)] = K + \lambda(1-p)t \end{aligned}$$

where we have used independence of N_1 and N_2

Example

- Suppose events occur in sequence. Events are of one of finitely many types. We want to calculate expected number of events for each type of event to occur at least once.
- e.g., collecting of discount coupons, collecting of 'cards', etc.
- Let us assume that, independently of everything else, each event that occurs is of Type-i with probability, p_i, $i = 1, \cdots, K$.
- This is like rolling a dice repeatedly and waiting for each outcome to occur at least once. (Outcomes are not equally likely!)

- Let N_i denote the number of events till the first occurrence of Type-i event.
- Then N_i is geometric with parameter p_i.
  But N_i need not be independent.
- What we want to EN where $N = \max_{1 \leq i \leq K} N_i$.
- Dealing with $\max$ of non-independent random variables is difficult.

- We can make the problem easier to solve by making more assumptions.
- We assume that the events (e.g., collection of coupons) occurs over time as a Poisson process, $N(t)$, with rate $\lambda = 1$.
- We would say that an event of this process is Type-i with probability p_i.
- Let $N_j(t)$ denote the process of Type-j events here. (It is a Poisson process with rate p_j (because $\lambda = 1$)
- The processes, $N_j(t), j = 1, \cdots, K$ are independent.

- Let X_j denote the time of the first event of the j^{th} process.
- Let $X = \max_{1 \leq j \leq K} X_j$
- Now X_j are independent exponential random variables.

$$\begin{aligned} Pr[X < t] &= Pr[X_j < t, \ j = 1, \cdots, K] \\ &= \prod_{j=1}^{K} \left(1 - e^{-p_j t}\right) \end{aligned}$$

- Hence we get

$$E[X] = \int_0^\infty Pr[X > t] \, dt = \int_0^\infty \left[1 - \prod_{j=1}^{K} \left(1 - e^{-p_j t}\right)\right] \, dt$$

- Is the problem solved?

- X is the time by which at least one event of each type has occurred.
- But we need the expected number of events by that time
- Let N denote the number of events by time X. Then

$$X = \sum_{i=1}^{N} T_i, \quad T_i \textit{ iid} \text{ exponential with mean } 1$$

- Hence, by Wald's identity, $E[X] = E[N]\ E[T_i] = E[N]$
- This completes the solution of the problem.

- There is an interesting generalization of this scenario.
- Events are of different types
- The type of an event can depend on the time of occurrence but it is independent of everything else.
- Suppose an event occurring at time t is Type-i with probability $p_i(t)$.
- $p_i(t) \geq 0, \ \forall i, t$ and $\sum_{i=1}^{K} p_i(t) = 1, \ \forall t$
- $N_i(t)$ is the number of Type-i events till t

Theorem; Then, at any t, $N_i(t), i = 1, \cdots K$ are independent Poisson random variables with

$$E[N_i(t)] = \lambda \int_0^t p_i(s) \, ds$$

The earlier case corresponds to $p_i(s) = p_i, \forall s$.

Example: Tracking infections

- We use a simple model
- Individuals get infected as a Poisson process with rate λ
- The infection has an incubation time – Time between getting infected and showing symptoms.
- We assume the incubation time is a random variable with known distribution function G
  An individual infected at s would show symptoms by t with probability $G(t-s)$

- The incubation time has a know distribution function, G.
- The incubation times of different infected individuals are iid
- Define
  - $N(t)$ – total number infected till t
  - $N_1(t)$ – number showing symptoms by t
  - $N_2(t)$ – infected by t but not showing symptoms
- We can observe (or estimate through sampling), the value of $N_1(t)$.
- The question is, how do we estimate $N_2(t)$

- We take t as current time and treat it as fixed
- Define two types of events.
  - An event occurring at s is Typ-1 with probability $G(t-s)$
  - It is Typ-2 with probability $1 - G(t-s)$
- Then, Typ-1 individuals are those showing symptoms by t
- From our theorem,

$$E[N_1(t)] = \lambda \int_0^t G(t-s)\, ds = \lambda \int_0^t G(y)\, dy$$

$$E[N_2(t)] = \lambda \int_0^t (1 - G(t-s))\, ds = \lambda \int_0^t (1 - G(y))\, dy$$

- Suppose we have n_1 people showing symptoms at t
- We can approximate

$$n_1 \approx E[N_1(t)] = \lambda \int_0^t G(y)\, dy$$

- Hence we can estimate

$$\hat{\lambda} = \frac{n_1}{\int_0^t G(y)\, dy}$$

- Using this we can approximate

$$E[N_2(t)] \approx \hat{\lambda} \int_0^t (1 - G(y))\, dy$$

Extensions: Non-homogeneous Poisson process

- The Poisson process we considered is called homogeneous because the rate is constant.
- For a non-homogeneous Poisson process the rate can be changing with time.
- But we can still use a definition similar to definition 2

$$Pr[N(t+h) - N(t) = 1] = \lambda(t)h + o(h)$$

- We still stipulate independent increments though we cannot have stationary increments now
- One can show that $N(t+s) - N(t)$ is Poisson with parameter $m(t+s) - m(t)$ where $m(\tau) = \int_0^\tau \lambda(s)\, ds$

Extensions: Compound Poisson process

- Suppose Y_i are iid and ind of $N(t)$. Then

$$X(t) = \sum_{i=1}^{N(t)} Y_i$$

  is called a compound Poisson process
- Customers arrive as a Poisson process. Money spent by customers are iid rv. Then revenue process is compound Poisson.
- By Wald's identity, $E[X(t)] = E[N(t)]E[Y_1] = \lambda t \; E[Y_1]$.

$$\begin{aligned} \mathsf{Var}(X(t)) &= E[N(t)]\mathsf{Var}(Y_1) + \mathsf{Var}(N(t))(E[Y_1])^2 \\ &= \lambda t \left(\mathsf{Var}(Y_1) + (E[Y_1])^2\right) = \lambda t E[Y_1^2] \end{aligned}$$