Poisson Process, Spike Train and All That

Zoran Nenadic
Division of Engineering and Applied Science
California Institute of Technology
Pasadena, CA 91125
1 Counting Process

Let Fig. 1 be a graph of the customers who are entering a bank. Every time a customer comes, the counter is increased by one. The time of the arrival of $i$-th customer is $t_i$. Since the customers are coming at random, the sequence $\{t_1, t_2, \ldots, t_m\}$, denoted shortly by $\{t_i\}$, is a random sequence. Also, the number of customers who came in the interval $(t_0, t]$ is a random variable (process). Such a process is right continuous, as indicated by the graph in Fig. 1.

As it is often case in the theory of stochastic processes, we assume that the index set, i.e. the set where $\{t_i\}$ is taking values from, is $T = [0, \infty)$. Therefore, we have a sequence of non-negative random variables

$$0 \leq t_0 < t_1 < t_2 < \cdots < t_m \to \infty \quad \text{as} \quad m \to \infty.$$  

WLOG\(^1\) let $t_0 = 0$ and $N_0 = 0$, then

$$N_t = \max\{n, t_n \leq t\}, \quad T = [0, \infty),$$

is called a point process (counting process), and is denoted shortly by $\{N_t, t \geq 0\}$.

---

\(^1\)Without loss of generality
Let $T_n \triangleq t_n - t_{n-1}$ be inter-arrival time, then the sequence of inter-arrival times $\{T_n, n \geq 1\}$ is another stochastic process.

Special case is when $\{T_n, n \geq 1\}$ is a sequence of i.i.d.\(^2\) random variables, then the sequence $\{t_n\}$ is called a **renewal process**. $\{N_t, t \geq 0\}$ is the associated renewal point process, sometimes also called renewal process. Also, keep in mind that $t_n = T_1 + T_2 + \cdots + T_n$.

**Definition** (Poisson process) A point process $\{N_t, t \geq 0\}$ is called a **Poisson process** if $N_0 = 0$ and $\{N_t\}$ satisfies the following conditions

1. its increments are stationary and its non-overlapping increments are independent

2. $P(N_{t+\Delta t} - N_t = 1) = \lambda \Delta t + o(\Delta t)$

3. $P(N_{t+\Delta t} - N_t \geq 2) = o(\Delta t)$

**Remarks**

- $\{N_t, t \in T\}; \ t, s \in T; \ t > s; \ N_t - N_s$ - is the increment of stochastic process $N_t$.

- $N_{t+\Delta t} - N_t$ is the number of new arrivals during $(t, t + \Delta t]$.

- $\lambda = \text{const} > 0$ and $o(\Delta t)$ is understood as $\frac{o(\Delta t)}{\Delta t} \to 0$ when $\Delta t \to 0$.

The Poisson process defined above is also known as **homogeneous Poisson process**. In general $\lambda$ can be a time dependent function $\lambda(t)$, in which case we are dealing with inhomogeneous Poisson process. Finally, $\lambda$ itself can be a realization of stochastic process $\lambda(t, \omega)$, in which case we have so-called **doubly stochastic Poisson process**.

In any case, the parameter $\lambda$ of a Poisson process is called the **rate** and sometimes the **intensity** of the process. Its dimension is $[\text{events}] / [\text{time}]$ (e.g. spikes/sec in neuroscience).

**Theorem** Let $\{N_t, t \geq 0\}$ be a Poisson process, then

$$P(N_t = k) = \frac{(\lambda t)^k}{k!} e^{-\lambda t} \quad k = 0, 1, \cdots$$  \hspace{1cm} (1)

\(^2\)Independent identically distributed
The expression on the left hand side of (1) represents the probability of \( k \) arrivals in the interval \((0, t]\).

**Proof** A generating function of a discrete random variable \( X \) is defined via the following z-transform (recall that the moment generating function of a continuous random variable is defined through Laplace transform):

\[
G_X(z) = E[z^X] = \sum_{i=0}^{\infty} z^i p_i,
\]

where \( p_i = P(X = i) \). Let us assume that \( X \) is a Poisson random variable with parameter \( \mu \), then

\[
P(X = i) = \frac{\mu^i}{i!} e^{-\mu} \quad i = 0, 1, 2, \ldots
\]

and

\[
G_X(z) = \sum_{i=0}^{\infty} z^i \frac{\mu^i}{i!} e^{-\mu} = e^{\mu(z-1)}.
\]

(2)

Going back to Poisson process, define the generating function as

\[
G_t(z) \triangleq E[z^{N_t}]
\]

Then we can write

\[
G_{t+\Delta t}(z) = E[z^{N_t+\Delta t}] = E[z^{N_t+N_{t+\Delta t}-N_t}] = E[z^{N_t}]E[z^{N_{t+\Delta t}-N_t}]
\]

\[
= G_t(z) [(1 - \lambda \Delta t + o(\Delta t)) z^0 + (\lambda \Delta t + o(\Delta t)) z^1 + o(\Delta t)(z^2 + \cdots)]
\]

Furthermore

\[
\frac{G_{t+\Delta t} - G_t(z)}{\Delta t} = G_t(z) \left[-\lambda + \frac{o(\Delta t)}{\Delta t} + (\lambda + \frac{o(\Delta t)}{\Delta t}) z + \frac{o(\Delta t)}{\Delta t} (z^2 + \cdots)\right]
\]

\[
\Rightarrow \lim_{\Delta t \to 0} \frac{G_{t+\Delta t} - G_t(z)}{\Delta t} = G_t(z) [-\lambda + \lambda z]
\]

\[
\Rightarrow \frac{dG_t(z)}{dt} = G_t(z) \lambda (z - 1)
\]

\[
\Rightarrow \log G_t(z) - \log G_0(z) = \lambda t (z - 1)
\]

\[
\Rightarrow G_t(z) = e^{\lambda t(z-1)}
\]
Comparing this result to (2) we conclude that $N_t$ is a Poisson random variable with parameter $\lambda t$. ■

**Theorem** If $\{N_t, t \geq 0\}$ is a Poisson process and $T_n$ is the inter-arrival time between the $n$-th and $(n-1)$-th events, then $\{T_n, n \geq 1\}$ is a sequence of i.i.d. random variables with exponential distribution, with parameter $\lambda$.

**Proof**

\[ P(T_1 > t) = P(N_t = 0) = e^{-\lambda t} \Rightarrow T_1 - \text{exponential} \]

Need to show that $T_1$ and $T_2$ are independent and $T_2$ is also exponential.

\[ P(T_2 > t \mid T_1 \in (s-\delta, s+\delta]) = \frac{P(T_2 > t, T_1 \in (s-\delta, s+\delta])}{P(T_1 \in (s-\delta, s+\delta])} \tag{3} \]

The event $\{T_2 > t, T_1 \in (s-\delta, s+\delta]\}$ is a subset of the event described by Fig. 2, i.e.

\[ P(T_2 > t, T_1 \in (s-\delta, s+\delta]] \leq P(\underbrace{N_{s-\delta} = 0, N_{s+\delta} - N_{s-\delta} = 1}, \underbrace{N_{s+t-\delta} - N_{s+\delta} = 0}) \]

no arrivals one arrival no arrivals

\[ P(T_2 > t, T_1 \in (s-\delta, s+\delta]) \leq P(T_1 \in (s-\delta, s+\delta]) P(N_{s+t-\delta} - N_{s+\delta} = 0) \]

\[ \Rightarrow P(T_2 > t, T_1 \in (s-\delta, s+\delta]) \leq P(T_1 \in (s-\delta, s+\delta]) e^{-\lambda(t-2\delta)} \]

From (3) \[ P(T_2 > t \mid T_1 \in (s-\delta, s+\delta]) \leq e^{-\lambda(t-2\delta)} \tag{4} \]
Similarly, the event described by Fig. 3 is a subset of the event \( \{T_2 > t, T_1 \in (s - \delta, s + \delta]\} \), therefore

\[
P(N_{s-\delta} = 0, N_{s+\delta} - N_{s-\delta} = 1, N_{s+t+\delta} - N_{s+\delta} = 0) \leq P(T_2 > t, T_1 \in (s - \delta, s + \delta])
\]

\[
\Rightarrow P(T_1 \in (s - \delta, s + \delta]) P(N_{s+t+\delta} - N_{s+\delta} = 0) \leq P(T_2 > t, T_1 \in (s - \delta, s + \delta])
\]

\[
\Rightarrow P(T_1 \in (s - \delta, s + \delta]) e^{-\lambda t} \leq P(T_2 > t, T_1 \in (s - \delta, s + \delta])
\]

From (3) \( \Rightarrow \)

\[
P(T_2 > t \mid T_1 \in (s - \delta, s + \delta]) \geq e^{-\lambda t}
\]

From (4) and (5), using squeeze theorem \((\delta \to 0)\), it follows

\[
P(T_2 > t \mid T_1 = s) = e^{-\lambda t} \quad \Rightarrow \quad f_{T_2 \mid T_1 = s}(t \mid s) = \lambda e^{-\lambda t}
\]

Therefore, \( T_2 \) is independent of \( T_1 \), and \( T_2 \) is exponentially distributed random variable. ■

**Theorem**

1. \( E[N_t] = \lambda t \)
2. \( Var[N_t] = \lambda t \)

**Proof** Recall that \( G_t(z) = E[z^{N_t}] \), then

\[
\left[ \frac{dG_t(z)}{dz} \right]_{z=1} = E \left[ N_t z^{N_t-1} \right]_{z=1} = E[N_t]
\]

\[
\Rightarrow E[N_t] = \left[ \lambda t e^{\lambda t(z-1)} \right]_{z=1} = \lambda t
\]
Likewise
\[
\left[ \frac{d^2 G_t(z)}{dz^2} \right]_{z=1} = E[N_t(N_t - 1)]
\]
\[\Rightarrow E[N_t^2] = \left[ (\lambda t)^2 e^{\lambda t(z-1)} \right]_{z=1} + E[N_t] = (\lambda t)^2 + \lambda t
\]
\[\Rightarrow Var[N_t] = (\lambda t)^2 + \lambda t - (\lambda t)^2 = \lambda t \quad \blacksquare
\]

**Theorem** (Conditioning on the number of arrivals) Given that in the interval \((0, T]\) the number of arrivals is \(N_T = n\), the \(n\) arrival times are independent and uniformly distributed on \([0, T]\).

**Proof** Independence of arrival times \(t_1, t_2\) etc. directly follows from independence of non-overlapping increments. In particular let \(t_1\) and \(t_2\) be arrival times of first and second event, then

\[
P(t_1 \in (0, s], t_2 \in (s, t]) = P(N_s = 1, N_t - N_s = 1) = P(N_s = 1) P(N_t - N_s = 1|N_s = 1) = P(t_1 \in (0, s]) P(t_2 \in (s, t])
\]

Suppose that we know exactly one event happened in the interval \((0, T]\), and suppose the interval is partitioned into \(M\) segments of length \(\Delta t\), as shown in Fig. 1. Let \(p_i\) be the probability of event happening in the \(i\)-th bin, then \(\sum_{i=1}^{M} p_i = 1\). From the definition of Poisson process it follows that \(p_i \propto \lambda \Delta t\), say \(p_i = C(\lambda \Delta t + o(\Delta t))\). The constant \(C\) is determined from

\[
\sum_{i=1}^{M} C(\lambda \Delta t + o(\Delta t)) = 1 \Rightarrow C = \frac{1}{\lambda M \Delta t + M o(\Delta t)} = \frac{1}{T(\lambda + o(\Delta t) \Delta t)}
\]

Let \(t_1\) be a random variable corresponding to the time of arrival, then the **probability density function** (pdf) of \(t_1\) can be defined as

\[
f_{t_1}(t) = \lim_{\Delta t \to 0} \frac{p_i}{\Delta t} = \frac{1}{T} \quad \forall i = 1, 2, \cdots, M \quad \text{where} \quad t = i \Delta t.
\]
Therefore, $t_1$ is uniformly distributed on $[0, T]$. Let $t_1$ and $t_2$ be the arrival times of two events, and we know exactly two events happened on $(0, T]$. Also assume that $t_1$ and $t_2$ represent mere labels of events, not necessarily their order. Given that $t_1$ happened in $j$-th bin, the probability of $t_2$ occurring in any bin of size $\Delta t$ is proportional to the size of that bin, i.e. $p_i \propto \lambda \Delta t$, except for the $j$-th bin, where $p_j \propto o(\Delta t)$. By rendering the bin size infinitesimal, we notice that the probability $p_i$ remains constant over all but one bin, the bin in which $t_1$ occurred, where $p_j = 0$. But this set is a set of measure zero, so the cumulative sum over $p_i$ again gives rise to uniform distribution on $(0, T]$. 

**Question** What is the probability of observing $n$ events at instances $\tau_1$, $\tau_2$, $\cdots$, $\tau_n$ on the interval $[0, T]$?

Since arrival times $t_1$, $t_2$, $\cdots$, $t_n$ are continuous random variables, the answer is 0. However, we can calculate the associated pdf as

$$f_{t_1, t_2 \cdots t_n}(\tau_1, \tau_2, \cdots, \tau_n) = \lim_{dt \to 0} \frac{P(t_1 \in (\tau_1, \tau_1 + dt], \cdots, t_n \in (\tau_n, \tau_n + dt], N_T = n)}{dt^n}$$

where

$$P(t_1 \in (\tau_1, \tau_1 + dt], \cdots, t_n \in (\tau_n, \tau_n + dt], N_T = n) = P(t_1 \in (\tau_1, \tau_1 + dt], \cdots, t_n \in (\tau_n, \tau_n + dt] \mid N_T = n) P(N_T = n)$$

$$= \left( \frac{dt}{T} \right)^n \frac{(\lambda T)^n}{n!} e^{-\lambda T} = \frac{\lambda^n dt^n}{n!} e^{-\lambda T}$$

$$\Rightarrow f_{t_1, t_2 \cdots t_n}(\tau_1, \tau_2, \cdots, \tau_n) = \frac{\lambda^n}{n!} e^{-\lambda T}$$

**Question** What is the power spectrum of Poisson process?

It does not make sense to talk about the power spectrum of Poisson process, since it is not a stationary process. In particular the mean of Poisson process is

$$E[N_t] = \lambda t$$
and its autocorrelation function is

\[ R(t, s) \triangleq E[N_t N_s] \]

\[ R(t, s) \overset{t \geq s}{=} E[(N_t - N_s + N_s) N_s] = E[(N_t - N_s) N_s + N_s^2] \]

\[ = E[N_t - N_s]E[N_s] + E[N_s^2] = \lambda(t - s)s + \lambda^2 s^2 + \lambda s = \lambda^2 t s + \lambda s \]

\[ R(t, s) \overset{t \leq s}{=} E[(N_s - N_t + N_t) N_t] = E[(N_s - N_t) N_t + N_t^2] \]

\[ = E[N_s - N_t]E[N_t] + E[N_t^2] = \lambda(s - t)t + \lambda^2 t^2 + \lambda t = \lambda^2 t s + \lambda t \]

Since \( R(t, s) \neq R(t - s) \), we conclude that \( \{N_t, t \geq 0\} \) is not stationary (in weak sense), therefore it does not make sense to talk about its power spectrum. Let us define the following stochastic process (Fig. 5)

\[ S_t = \frac{dN_t}{dt} = \sum_i \delta(t - t_i) \quad \text{spike train} \quad (6) \]

The fundamental lemma says that if \( Y(t) = L\{X(t)\} \), where \( L \) is a linear operator, then

\[ E[Y(t)] = L\{E[X(t)]\} \]

Since differentiation is a linear operator we have

\[ E[S_t] = \frac{d(\lambda t)}{dt} = \lambda \]
Also, it can be shown using theory of linear operators that
\[
R_{SS}(t, s) = \frac{\partial}{\partial t} R_{NN}(t, s) = \begin{cases} 
\frac{\partial}{\partial t} \frac{\lambda^2}{2} t + \lambda & t > s \\
\frac{\partial}{\partial s} \frac{\lambda^2}{2} t & t < s
\end{cases}
\]

Heaviside function

Thus, \( S_t \) is WWS\(^3 \) stochastic process, and it makes sense to define the power spectrum of such a process as a Fourier transform of its autocorrelation function i.e.

\[
P_S(\omega) = \mathcal{F}\{R_{SS}(\tau)\} = \int_{-\infty}^{\infty} R_{SS}(\tau) e^{-j\omega \tau} d\omega = \lambda^2 + \lambda^2 \delta(t - s)
\]

Therefore, the spike train \( S_t = \sum_i \delta(t - t_i) \) of independent times \( t_i \) behaves almost as a white noise, since its power spectrum is flat for all frequencies, except for the spike at \( \omega = 0 \). The process \( S_t \) defined by (6) is a simple version of what is in engineering literature known as a shot noise.

**Definition** (Inhomogeneous Poisson process) A Poisson process with a non-constant rate \( \lambda = \lambda(t) \) is called inhomogeneous Poisson process. In this case we have

1. non-overlapping increments are independent (the stationarity is lost though).
2. \( P(N_{t+\Delta t} - N_t = 1) = \lambda(t) \Delta t + o(\Delta t) \)
3. \( P(N_{t+\Delta t} - N_t \geq 2) = o(\Delta t) \)

**Theorem** If \( \{N_t, t > 0\} \) is a Poisson process with the rate \( \lambda(t) \), then \( N_t \) is a Poisson random variable with parameter \( \mu = \int_0^t \lambda(\xi) d\xi \) i.e.

\[
P(N_t = k) = \frac{(\int_0^t \lambda(\xi) d\xi)^k}{k!} e^{-\int_0^t \lambda(\xi) d\xi} \quad (7)
\]

\(^3\)Wide (weak) sense stationary. A stochastic process \( X(t) \) is WSS if \( E[X(t)] = const \) and \( R_{XX}(t, s) = R_{XX}(t - s) \)
Proof The proof of this theorem is identical to that of homogeneous case except that \( \lambda \) is replaced by \( \lambda(t) \). In particular, one can easily get
\[
G_t(z) = e^{(z-1) \int_0^t \lambda(\xi) \, d\xi},
\]
from which (7) readily follows.

Theorem Let \( \{N_t, t > 0\} \) be an inhomogeneous Poisson process with the rate \( \lambda(t) \) and let \( t > s \geq 0 \), then
\[
P(N_t - N_s = k) = \left( \int_s^t \lambda(\xi) \, d\xi \right)^k \frac{e^{-\int_s^t \lambda(\xi) \, d\xi}}{k!}
\]
The application of this theorem stems from the fact that we cannot use \( P(N_t - N_s = k) = P(N_{t-s} = k) \), since the increments are no longer stationary.

Proof
\[
G_t(z) = E[z^{N_t}] = E[z^{N_t - N_s + N_s}] = E[z^{N_t - N_s}] E[z^{N_s}] = E[z^{N_t - N_s}] G_s(z)
\]
\[
\Rightarrow E[z^{N_t - N_s}] = \frac{G_t(z)}{G_s(z)} = \frac{e^{(z-1) \int_s^t \lambda(\xi) \, d\xi}}{e^{(z-1) \int_s^t \lambda(\xi) \, d\xi}} = e^{(z-1) \int_s^t \lambda(\xi) \, d\xi}
\]
Thus, \( N_t - N_s \) is a Poisson random variable with parameter \( \mu = \int_s^t \lambda(\xi) \, d\xi \), and (9) easily follows.

Theorem
1. \( E[N_t] = \int_0^t \lambda(\xi) \, d\xi \)
2. \( Var[N_t] = \int_0^t \lambda(\xi) \, d\xi \)

Proof Recall that
\[
E[N_t] = \left[ \frac{dG_z(t)}{dz} \right]_{z=1} \quad \text{and} \quad E[N_t^2] = \left[ \frac{d^2G_z(t)}{dz^2} \right]_{z=1} + E[N_t]
\]
From (8) we have \( G_t(z) = e^{(z-1) \int_0^t \lambda(\xi) \, d\xi} \), and the two results follow after immediate calculations.
**Theorem** (Conditioning on the number of arrivals) Given that in the interval \((0, T]\) the number of arrivals is \(N_T = n\), the \(n\) arrival times are independently distributed on \([0, T]\) with the pdf \(\lambda(t)/ \int_0^T \lambda(\xi) d\xi\).

**Proof** The proof of this theorem is analogous to that of the homogeneous case. The probability of a single event happening at any of \(M\) bins (Fig. 1) is given by \(p_i = C(\lambda(i \Delta t) \Delta t + o(\Delta t))\), where \(i\) is the bin index. Given that exactly one event occurred in the interval \((0, T]\), we have

\[
\sum_{i=1}^{M} p_i = 1 \Rightarrow C = \frac{1}{\sum_{i=1}^{M} \lambda(i \Delta t) \Delta t + T \frac{d(\Delta t)}{\Delta t}}
\]

\[
f_{t_i}(t) = \lim_{\Delta t \to \infty} p_i \frac{\lambda(t)}{\int_0^T \lambda(\xi) d\xi} \quad \text{where} \quad t = i \Delta t.
\]

The argument for independence of two or more arrival times is identical to that of the homogeneous case.

**Question** What is the probability of observing \(n\) events at instances \(\tau_1, \tau_2, \cdots, \tau_n\) on the interval \([0, T]\)?

Since arrival times \(t_1, t_2, \cdots, t_n\) are continuous random variables, the answer is 0. However, we can calculate the associated pdf as

\[
f_{t_1, t_2 \cdots t_n}(\tau_1, \tau_2, \cdots, \tau_n) = \lim_{dt \to 0} \frac{P(t_1 \in (\tau_1, \tau_1 + dt], \cdots, t_n \in (\tau_n, \tau_n + dt], N_T = n)}{dt^n}
\]

where

\[
P(t_1 \in (\tau_1, \tau_1 + dt], \cdots, t_n \in (\tau_n, \tau_n + dt], N_T = n) = \frac{\prod_{i=1}^{n} \int_{\tau_i}^{\tau_i + dt} \lambda(\tau) d\tau}{\int_0^T \lambda(\xi) d\xi} \frac{(\int_0^T \lambda(\xi) d\xi)^n}{n!} e^{-\int_0^T \lambda(\xi) d\xi}
\]

\[
\Rightarrow f_{t_1, t_2 \cdots t_n}(\tau_1, \tau_2, \cdots, \tau_n) = \frac{\prod_{i=1}^{n} \lambda(\tau_i)}{n!} e^{-\int_0^T \lambda(\xi) d\xi}
\]
Figure 6: Realization of a point process using conditioning on the number of arrivals. (Top) Ten different sample paths of the same point process shown as raster plots. (Bottom) The histogram of inter-arrival times, showing the exponential trend.
Figure 7: Realization of a point process using method of infinitesimal increments. (Top) Ten different sample paths of the same point process shown as raster plots. (Bottom) The histogram of inter-arrival times, showing the exponential trend.
Figure 8: Realization of a point process using method of independent inter-arrival times. (Top) Ten different sample paths of the same point process shown as raster plots. (Bottom) The histogram of inter-arrival times, showing the exponential trend.