Poisson Process

Poisson distribution is often used to model count data.
e.g. in phone calls coming into a center given a time interval

(o) arrivals on a carafe

(3) # mutant sites of a newborn

One of the basic assumptions on which Poisson is built is that, for small time intervals, the
prob of an arrival is proportional to the length of waiting time.

Let \( N(t) \) be the # of events of some specific kind from time \( 0 \) to \( t \).
\( N(t) \) is a stochastic process indexed by time.

Consider \( \Delta N(t) : t > 0 \)
define \( N(0) = 0 \)

Definition. A Poisson process is defined as a counting process that satisfies the
following conditions/assumptions

1. **Stationary** \( N(t+s) - N(t) \sim N(s) \)
   have the same distribution
   # events occurred in a time interval only depends on interval length.
   is independent of the time that the interval starts

2. **Independent Increment**
   Let \( 0 < t_1 < t_2 < \ldots \), we assume \( N(t_1) \mid N(t_1) - N(t_1 - t) \) \( \mid N(t_1 - t) \) \( \mid \ldots \)
   mutual independence of non-overlapping intervals

3. **For very small \( \Delta t \), \( P(N(\Delta t) = 0) = \lambda \Delta t + o(\Delta t) \) for some \( \lambda > 0 \), and**
   \( P(N(\Delta t) = 1) = \lambda \Delta t \)
   This says that for small intervals, the prob of an arrival/event is proportional
   to the length of waiting time, the prob of two or more events is negligible.
   Here \( o(x) \) stands for functions such that \( \frac{g(x)}{x} \to 0 \) as \( x \to 0 \).

E.g. Let \( \Delta t = \frac{t}{n} \)

\[ \frac{\Delta t}{\Delta \Delta t} = \left( \frac{\Delta t}{\Delta t} \right)^2 \to 0 \text{ as } \Delta t \to 0 \]
Claim: \( P(N(t)=0) = e^{-\lambda t} \)

**Proof** Break \((0,t)\) into \(n\) intervals

\[
P(N(t)=0) = \lim_{n \to \infty} \left[ \prod_{k=1}^{n} P(N(k^{th} t)-N(k^{th} t)=0) \right] = \lim_{n \to \infty} \left[ \prod_{k=1}^{n} P(N(k^{th} t)=0) \right]
\]

Stationary

\[
\lim_{n \to \infty} \left( \frac{n}{n} \right) = 1
\]

\[
\lim_{n \to \infty} \left( \frac{n}{n} \right) = e^{-\lambda t}
\]

Claim: \( P(N(t)=k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!} \), i.e., \( N(t) \sim \text{Poisson}(\lambda t) \)

**Proof** Back to the partition of \((0,t)\)

Define events \( A = \{N(t)=k\}\) and \( B = \{N(t)=k\} \) and at least one interval has \( \geq 2\) events.

\[
P(A) = \lim_{n \to \infty} \left( \frac{n}{k} \right) \left( \frac{P(N(k^{th} t)=1)}{1-P(N(k^{th} t) \geq 2)} \right) \left( \frac{P(N(k^{th} t)=0)}{1-P(N(k^{th} t) \geq 2)} \right)^n = \lim_{n \to \infty} \left( \frac{n}{k} \right) \left[ \frac{\lambda t}{n} + \frac{t}{n} \right] \left( 1- \frac{\lambda t}{n} + \frac{t}{n} \right) \left( 1- \frac{\lambda t}{n} + \frac{t}{n} \right)^n = \lim_{n \to \infty} \left( \frac{n}{k} \right) \left( \frac{\lambda t}{n} \right)^k \left( 1- \frac{\lambda t}{n} + \frac{t}{n} \right)^n = \frac{(\lambda t)^k e^{-\lambda t}}{k!}
\]

\( \leq P(B) = P(\text{at least one interval has } \geq 2 \text{ events}) \)

\[
\leq \lim_{n \to \infty} \left( \frac{n}{k} \right) \left[ \frac{\lambda t}{n} + \frac{t}{n} \right] \left( 1- \frac{\lambda t}{n} + \frac{t}{n} \right) \left( 1- \frac{\lambda t}{n} + \frac{t}{n} \right)^n = \frac{(\lambda t)^k e^{-\lambda t}}{k!}
\]

So \( P(B) = 0 \)

\[
P(N(t)=k) = P(A) + P(B) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}
\]
Note: There are alternative but equivalent definitions for Poisson Process.

For example, we can use the following definition:

1. Mutually independence of non-overlapping intervals
2. \( N(t) - N(s) \sim \text{Poisson}(\lambda(t-s)) \) for \( s < t \)

**Proof.** Based on the proof of the previous page, \( 2 \Rightarrow 1 \)

1. \( N(s) = N(S) + N(t) \sim \text{Poisson}(\lambda(s)) = \text{Poisson}(\lambda s) \)

2. \( N(t) = N(t) \sim \text{Poisson}(\lambda(t-s)) = \text{Poisson}(\lambda t - \lambda s) \)

3. \( N(t) - N(s) \sim \text{Poisson}(\lambda(t-s)) = \text{Poisson}(\lambda t - \lambda s) \)

So \( N(t)-N(s) \sim \text{Poisson}(\lambda(t-s)) = \text{Poisson}(\lambda t - \lambda s) \)

**Waiting time.**

Let \( T \) denote the waiting time for the first event.

- \( P(T > t) = P(N(t) = 0) = \frac{(at)^0 e^{-at}}{0!} = e^{-at} \)
- \( F_T(t) = P(T < t) = 1 - e^{-at} \)
- \( f_T(t) = a e^{-at} \sim \text{Exp}(\frac{1}{a}) \) where \( \lambda \) is rate

**Conditional Distribution.** Suppose \( N(t) = n \). What is the probability of \( N(s) = k \) where \( 0 < s < t \)?

- \( P(N(s) = k | N(t) = n) = \frac{P(N(s) = k, N(t) = n)}{P(N(t) = n)} = \frac{P(N(s) = k, N(t)-N(s) = n-k)}{P(N(t) = n)} \)
  - \( = \frac{P(N(s) = k)}{P(N(t) = n)} \cdot \frac{P(N(t)-N(s) = n-k)}{P(N(t)-N(s) = n-k)} \)
  - \( = \frac{n! k!}{(n-k)! k!} \cdot \frac{(at-k)^{n-k} e^{-at}}{(n-k)!} \)

So \( N(s) | N(t) = n \sim \text{Binomial}(n, \frac{\lambda}{t}) \)
E.g. Taxi cabs arrive at a taxi stand according to Poisson Process with rate \( \lambda = 3 \) per hour.

Let \( N(t) \) = # arrivals by time \( t \) \( \sim \) \( \frac{(\lambda t)^k e^{-\lambda t}}{k!} \)

\[
P(N_0=0) = \frac{0^k e^{-3}}{0!} = e^{-3} \approx \frac{1}{30}
\]

\[
P(\text{exactly 3 cabs in 1 hr and exactly none in next 3 hrs and exactly 1 in next hr}) = P(N(1) = 3, N(4) - N(1) = 0, N(5) - N(4) = 1)
\]

\[
= P(N(1) = 3) P(N(4) - N(1) = 0) P(N(5) - N(4) = 1)
\]

\[
= \frac{3! e^{-3} \frac{3^3}{3!}}{3!} \cdot \frac{e^{-3}}{1!} = \frac{27}{2} e^{-15}
\]