OM9103: Stochastic Process

Lectures 3&4:
Poisson Process

A. Poisson Process

A.1. <u>*Definition*</u>: A stochastic process $\{N(t),\ t \geq 0\}$is called *a counting process* if $N(t)$ is a random variable representing the number of events that have occurred up to time t.

(i). $N(t) \geq 0$ for all $t \geq 0$.
(ii). $N(t) \in \mathbf{N}$ for all $t \geq 0$.
(iii). If $s < t$, $N(s) \leq N(t)$.
(iv). For $s < t$, then $N(t) - N(s)$ is equal to the number of events that have occurred in the interval $(s, t]$.

Examples:

- Count the number of cars that pass a certain intersection;
- Count the number of customers that arrive at a bank teller, supermarket counter, etc.;
- Bernoulli process
- "Limit" of the Bernoulli process: Poisson process

A.2. <u>*Proposition*</u>: (*Characterization 1*): $\{N(t),\ t \geq 0\}$is a Poisson process with rate λ if and only if

(i). $N(0) = 0$;
(ii). The process has independent increments;
(iii). The number of events in any interval of length t is Poisson distributed with mean λt: that is, for , for all $t, s \geq 0$,

$$\Pr(N(t+s) - N(s) = n) = e^{-\lambda t}\frac{(\lambda t)^n}{n!}, n = 0,1,\cdots.$$

The third condition implies that a Poisson process has stationary increments, and $E(N(t)) = \lambda t$. This is why λ is called *the rate of the Poisson process*.

A.3. <u>*Proposition*</u>: (*Characterization 2*): $\{N(t),\ t \geq 0\}$is a Poisson process with rate λ if and only if

(i). $N(0)=0$;
(ii). The process has stationary and independent increments;
(iii). $\Pr(N(h) = 1) = \lambda h + o(h)$;
(iv). $\Pr(N(h) \geq 2) = o(h)$.

The 3rd and 4th conditions can alternatively be stated as:

$$\left.\frac{d\Pr(N(h)=1)}{dh}\right|_{h=0} = \lambda; \qquad \left.\frac{d\Pr(N(h)\geq 2)}{dh}\right|_{h=0} = 0.$$

Ross proves that Characterization 2 implies characterization 1 (p.61-63). In addition, Characterization 1 implies characterization 2:

$$\frac{d\,\Pr(N(h)=1)}{dh}=\frac{d}{dh}\left(\lambda he^{-\lambda h}\right)=\lambda e^{-\lambda h}-\lambda^2 he^{-\lambda h};$$
$$\frac{d\,\Pr(N(h)\geq 2)}{dh}=\frac{d}{dh}\left(1-e^{-\lambda h}-\lambda he^{-\lambda h}\right)=\lambda^2 he^{-\lambda h}.$$

The result now follows by substituting $h = 0$.

A.4 *The Conditional Distribution of the Number of Arrivals*

Let $\{N(t), t \geq 0\}$ be a Poisson process with rate λ. Consider the probability distribution of $N(s)$ given that $N(t) = n$, where $t > s$. In other words, given n events in $[0, t]$, what is the distribution of the number of events in the first s time units of that interval?

$$\Pr(N(s)=k \mid N(t)=n)=\frac{\Pr(N(s)=k, N(t)=n)}{\Pr(N(t)=n)}$$
$$=\frac{\Pr(N(s)=k, N(t)-N(s)=n-k)}{\Pr(N(t)=n)}=\frac{\Pr(N(s)=k)\Pr(N(t)-N(s)=n-k)}{\Pr(N(t)=n)}$$
$$=\frac{e^{-\lambda s}(\lambda s)^k}{k!}\cdot\frac{e^{-\lambda(t-s)}(\lambda(t-s))^{n-k}}{(n-k)!}\cdot\frac{n!}{e^{-\lambda t}(\lambda t)^n}=\binom{n}{k}\left(\frac{s}{t}\right)^k\left(1-\frac{s}{t}\right)^{n-k}$$

That is, $N(s)|N(t) = n$ has a binomial distribution with parameters $(n, s/t)$. In particular, the probability that a given event took place in $[0, s]$ given that there are n events in $[0, t]$ is s/t.

B. Interarrival and Waiting Time

B.1. Interarraival Times

Based on typical applications of a Poisson process, an *event is* often also called an *arrival*. Given a Poisson process $\{N(t), t \geq 0\}$, define the *interarrival times* of consecutive events by

$$X_n = \min\{t\text{: } N(t)\geq n\} - \min\{t\text{: } N(t)\geq n\text{-}1\}$$

which is the time between $(n – 1)$st and the nth event.

It is easy to see that

$$\Pr(X_1 > t) = \Pr(N(t) = 0) = e^{-\lambda t}$$

That is, X_1 is exponentially distributed with rate λ.

To obtain the distribution for X_2, we condition it on X_1:

$\Pr(X_2 > t \mid X_1 = s) = \Pr(0 \text{ events in } (s, s + t] \mid X_1 = s)$
$= \Pr(0 \text{ events in } (s, s + t])$ (by independent increments)
$= e^{-\lambda t}$ (by stationary increments)

Hence X_2 is exponentially distributed with parameter λ as well, and furthermore, independent of X_1. In general, we have the following result:

Proposition: For any Poisson Process $\{N(t),\ t \geq 0\}$ with a parameter (rate) λ, the corresponding interarrival times X_n, $n = 1, 2, \ldots$, are independent identically distributed exponential random variables having mean $1/\lambda$.

B.2. Waiting Times

We can actually use the above proposition to derive any Poisson process. Let $X_1, X_2, \ldots$ be i.i.d. exponentially distributed random variables with parameter λ. Consider the arrival time of the n^{th} event:

$$S_n = \sum_{i=1}^{n} X_i,$$

which follows Gamma (or Erlang) distributed with parameters n and λ (hint: using the moment generating functions to prove it). Its probability density function is

$$f(t) = \frac{\lambda e^{-\lambda t}(\lambda t)^{n-1}}{(n-1)!}, \text{ for all } t \geq 0.$$

Proposition: Given that $X_1, X_2, \ldots$ be i.i.d. exponentially distributed random variables with parameter λ. Then the stochastic process defined by

$$N(t) = \max\left\{ n : S_n = \sum_{i=1}^{n} X_i \leq t \right\}$$

is a Poisson process with rate λ.

- In line with Characterization 1, it suffices to validate the third property:

$$\Pr(N(t) = n) = \Pr(S_n \leq t, S_{n+1} > t) = \int_0^t \Pr(S_{n+1} > t \mid S_n = s) \frac{\lambda e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!} ds$$

$$= \int_0^t \Pr(X_{n+1} > t - s) \frac{\lambda e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!} ds = \int_0^t e^{-\lambda(t-s)} \frac{\lambda e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!} ds$$

$$= \int_0^t \frac{e^{-\lambda t}\lambda^n s^{n-1}}{(n-1)!} ds = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$$

A useful relationship between the counting process and its interarrival/waiting times is given by the equivalence of the following two events:

$$N(t) \geq n \Leftrightarrow S_n \equiv \sum_{i=1}^{n} X_i \leq t$$

C. Conditional Distribution of Arrival Times

Consider that $N(t) = 1$ for some t. What is the distribution of the time at which that single event took place? Since a Poisson process has stationary and independent increments, we should expect that each interval in [0, 1] of equal length should have the same probability of containing the event. This implies that the time of the event (arrival time) is uniformly distributed in [0, t].

$$\Pr(X_1 < s \mid N(t)=1) = \frac{P(X_1 < s, N(t)=1)}{P(N(t)=1)} = \frac{\Pr(1 \text{ events in } [0, s), 0 \text{ event in } [s, t))}{P(N(t)=1)}$$

$$= \frac{\Pr(1 \text{ events in } [0, s)) \cdot \Pr(0 \text{ event in } [s, t))}{P(N(t)=1)} = \frac{\lambda s e^{-\lambda s} \cdot e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} = \frac{s}{t}.$$

Now consider an arbitrary value of n, i.e., we again condition on $N(t) = n$ for some t. What is the distribution of the time at which the *first* of these n events took place?

$$\Pr(X_1 > s \mid N(t)=n) = \Pr(N(s)=0 \mid N(t)=n) = \left(1-\frac{s}{t}\right)^n$$

Note that this is also the distribution of the 1st order statistic of a random sample of size n from the uniform distribution on [0, t]! This is a very important insight.

What is the distribution of the time at which the second of the n events took place?

$$\Pr(S_2 > s \mid N(t)=n) = \Pr(N(s) \le 1 \mid N(t)=n) = n\frac{s}{t}\left(1-\frac{s}{t}\right)^{n-1} + \left(1-\frac{s}{t}\right)^n$$

Note that this is also the distribution of the 2nd order statistic of a random sample of size n from the uniform distribution on [0, t]. In general, we have the following result:

<u>*Theorem 2.3.1*</u>: Given that $N(t) = n$, the arrival times $S_1, \ldots, S_n$ have the same distribution as the order statistics corresponding to n independent random variables uniformly distributed on the interval (0, t).

<u>Proof</u>: Refer to the text (p.67).

Put differently, the unordered arrival times given that exactly n events took place in [0, t] are i.i.d. according to the uniform distribution on [0, t].

<u>*Corollary* (Example 2.3(a)</u>: (**Expected Total Waiting Times**). Given any Poisson process $\{N(t): t \ge 0\}$ with rate λ, let S_i be the arrival time of the ith arrival. Then, the expected value of the total waiting times for all arrivals is given by:

$$E\left[\sum_{i=1}^{N(t)} (t - S_i)\right] = \frac{\lambda t^2}{2} \left(= \frac{t}{2} E(N(t))\right)$$

<u>Proof</u>: We will first find the conditional expected value given $N(t) = n$:

$$E\left[\sum_{i=1}^{N(t)} (t-S_i) \mid N(t)=n\right] = E\left[\sum_{i=1}^{n} (t-S_i) \mid N(t)=n\right] = nt - E\left[\sum_{i=1}^{n} S_i \mid N(t)=n\right].$$

Let $U_1, U_2, \ldots, U_n$ be a set of n independent uniform (0, t) random variable with the corresponding order statistics $U_{(1)}, U_{(2)}, \ldots, U_{(n)}$. Then according to Theorem 2.3.1, we have

$$E\left[\sum_{i=1}^{n} S_i \mid N(t)=n\right] = E\left[\sum_{i=1}^{n} U_{(i)}\right] = E\left[\sum_{i=1}^{n} U_i\right] = \frac{nt}{2}.$$

Hence it follows that

$$E\left[\sum_{i=1}^{N(t)}(t-S_i)\right]=E\left(E\left[\sum_{i=1}^{N(t)}(t-S_i)\mid N(t)=n\right]\right)=\frac{t}{2}E\left[N(t)\right]=\frac{\lambda t^2}{2}.$$

Inspection Paradox & Distribution of Age

Consider a piece of equipment, for instance, lightbulb or battery, whose lifetime can be approximated by an exponential distribution with parameter λ. As soon as the piece of equipment fails, it is replaced by a new one. Then the number of replacements up to time t is a Poisson process $\{N(t), t \geq 0\}$with rate λ. Now fix some (future) time t, and consider the piece of equipment that is in use at that point in time.

- What is the expected (total) lifetime of that item?
- What is the distribution of the lifetime of that item?

Write the lifetime of the item as follows:

$$X_{N(t)+1}=A(t)+Y(t)$$

where $A(t)$ is the age of the inspected item and $Y(t)$ is the remaining lifetime of the item. Clearly, $Y(t)$ has the exponential distribution, so

$$E(Y(t))=1/\lambda.$$

It is also easy to see that $A(t)$ is nonnegative, so $E(A(t)) \geq 0$. In addition, $A(t)$ is strictly positive with positive probability, so $E(A(t)) > 0$. Therefore,

$$E(X_{N(t)+1})=E(A(t))+E(Y(t))>1/\lambda.$$

This property is called ***the inspection paradox***. It says that the expected total lifetime of the item that is in use at some future time t is larger than the expected total lifetime of a typical item. Can you explain this intuitively?

We can actually derive the distribution of the age of the item in use at time t, $A(t)$:

$$\Pr(A(t) > x) = \Pr(0 \text{ replacement in } [t-x,\, t]) = \begin{cases} e^{-\lambda x} & x<t \\ 0 & otherwise \end{cases}$$

This is a truncated exponential distribution, with all mass corresponding to realizations $\geq t$ concentrated on t. This distribution is mixed with a mass at t:

$$\Pr(A(t) = t) = e^{-\lambda t}.$$

Hence the expected value of the age of the item in use at time t is:

$$E(A(t))=\int_0^\infty \Pr(A(t)>x)dx=\int_0^t e^{-\lambda x}dx=\frac{1-e^{-\lambda t}}{\lambda}.$$

Note that the expected total lifetime of the item that is in use at time t is equal to

$$E(X_{N(t)+1})=E(A(t))+E(Y(t))=\frac{1-e^{-\lambda t}}{\lambda}+\frac{1}{\lambda}=\frac{2-e^{-\lambda t}}{\lambda}$$

Asymptotically, this means that

$$\lim_{t\to\infty}E(X_{N(t)+1})=\frac{2}{\lambda}.$$

In other words, for t large, the expected lifetime of an inspected item is about twice as large as the expected lifetime of a typical item!

Note that, for large t the Poisson process looks stochastically the same whether going backwards or forwards in time. Thus, both the age as well as remaining lifetime at time t is (approximately) exponentially distributed.

D. Multiple Event Types and Filtered Process

D.1 Filtered Processes

Consider a Poisson process $\{N(t), t \geq 0\}$ with rate λ. The process counts events, which can be of one of two types. Let the probability that the i^{th} event is of type 1 be equal to p, independent of i and of the event times. Then the probability that the i^{th} event is of type 2 be equal to $(1 - p)$. It is evident that the newly derived process $\{N_k(t), t \geq 0\}$ that counts the events of type k only ($k = 1, 2$) is also a counting process. We call this process a **filtering** of the original process. Note that $N_k 0) = 0$. The questions are: Does this new process have independent and stationary increments?

Let us analyze the distribution of the random variable $N_1(t)$:

$$\Pr(N_1(t) = m) = \sum_{n=m}^{\infty} \Pr(N_1(t) = m \mid N(t) = n) \cdot \Pr(N(t) = n)$$

$$= \sum_{n=m}^{\infty} \binom{n}{m} p^m (1-p)^{n-m} \cdot e^{-\lambda t} \frac{(\lambda t)^n}{n!} = \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} p^m (1-p)^{n-m} \cdot e^{-\lambda t} \frac{(\lambda t)^n}{n!}$$

$$= \frac{(p\lambda t)^m}{m!} e^{-\lambda t} \sum_{n=m}^{\infty} \frac{((1-p)\lambda t)^{n-m}}{(n-m)!} = \frac{(p\lambda t)^m}{m!} e^{-\lambda t} e^{(1-p)\lambda t} = \frac{(p\lambda t)^m}{m!} e^{-p\lambda t}$$

Thus $N_1(t) \sim \text{Poisson}(p\lambda t)$.

Proposition: Given a Poisson process $\{N(t), t \geq 0\}$ with rate λ that has two types of events with probability occurrences of p and 1-p respectively. Then the derived filtered processes $\{N_k(t), t \geq 0\}$ are in fact two independent Poisson processes with rate $p\lambda$ (for $k = 1$)and $(1$-$p)\lambda$ (for $k = 2$) respectively.

Proof: We only need to prove the independence of the two processes:

$$\Pr(N_1(t) = n_1, N_2(t) = n_2)$$

$$= \Pr(N_1(t) = n_1, N_2(t) = n_2 \mid N(t) = n_1 + n_2) \cdot \Pr(N(t) = n_1 + n_2)$$

$$= \binom{n_1 + n_2}{n_1} p^{n_1} (1-p)^{n_2} \cdot e^{-(1-p)\lambda t} \frac{((1-p)\lambda t)^{n_1+n_2}}{(n_1 + n_2)!}$$

$$= e^{-p\lambda t} \frac{(p\lambda t)^{n_1}}{n_1!} \cdot e^{-(1-p)\lambda t} \frac{((1-p)\lambda t)^{n_2}}{n_2!}$$

which proves the independence of the two processes.

For more than two types of events, we have the following extension:

- Consider a Poisson process $\{N(t), t \geq 0\}$ with parameter λ.
- The process counts events, which can be one of m types.
- Let the probability that the i^{th} event is of type k be equal to p_k, independent of i and of the event times.

- Then the corresponding filtered process $\{N_k(t),\ t \geq 0\}$ is a Poisson process with parameter $p_k\lambda$.
- Furthermore, these underlying filtered Poisson processes are independent as well.

Similar to the process of ***filtering*** a Poisson process counting events of different types, we may ***aggregate*** Poisson processes.

Proposition (Aggregation of Independent Poisson Processes): If $\{N_j(t),\ t \geq 0\}$ are independent Poisson processes with parameters λ_j, then

$$\left\{ N(t) = \sum_{j=1}^{m} N_j(t), t \geq 0 \right\}$$

is a Poisson process with parameter $\sum_{j=1}^{m} \lambda_j$.

D.2 Multiple Event Types with Dynamic Classifications

Consider a Poisson process $\{N(t), t \geq 0\}$ with parameter λ. The process counts events, which can be one of m types. Let the probability that an event taking place at time s is of type k be equal to $P_k(s)$, independent of previous event times.

Proposition 2.3.2 (p.69): Let $\{N_k(t), t \geq 0\}$denote the process that counts the events of type k only. Then for any fixed t, the random variables $N_1(t),\ldots, N_m(t)$ are independent Poisson distributed having respective means $\lambda p_1 t, \ldots, \lambda p_m t$, where p_k is defined as follows:

$$p_k = \frac{1}{t}\int_0^t P_k(s)ds.$$

Is $\{N_k(t), t \geq 0\}$ also a Poisson process? Recall that if so, we should have the following:

$$E(N_k(t)) = \lambda_k t \quad \text{(for some } \lambda_k\text{)}.$$

By the previous result, we have

$$E(N_k(t)) = \lambda t \cdot \frac{1}{t}\int_0^t P_k(s)ds,$$

which requires that p_k is independent of t, which is equivalent to requiring,

$$P_k(s) = \frac{\lambda_k}{\lambda} \text{ for all } s.$$

Example 2.3(B): $M/G/\infty$ - Infinite Server Poisson Queue.

- Customer arrivals follows a Poisson process with rate λ, this is the same thing as requiring that the interarrival time follows an exponential distribution – this is the meaning of "M" in the expression;
- Upon arrival the customer is immediately served by one of the infinite number of possible servers (the source of "∞"); and
- The service times are assumed be independent with a common distribution G (the source of "G").

We are interested in the distribution of the number of customers that have completed their services (and left the system) (called **Type-I** customers) – denoted by $N_1(t)$, and the number of customers that are in service (called **Type-II** customers) – denoted by $N_2(t)$.

Note that if a customer enters at time s, $s \le t$, then it will be a type-I customer if the service times is less than $t - s$. Since the service time distribution is G, the probability of a type-I customer occurrence will be $P(s) = G(t - s)$. Then by the above proposition, the distribution of $N_1(t)$ is Poisson with mean:

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

Similarly,

$$E[N_2(t)] = \lambda \int_0^t \overline{G}(y)dy.$$

Furthermore, $N_1(t)$ and $N_2(t)$ are independent.

The following example has the identical structure as in the above example.

Example: (**Infectious Disease Spreading**) Consider a disease, spread by an insect. The incubation time of the disease is very large. That is, many people could be infected without being aware of it. Can we say something about the total number of people that are infected with disease?

Suppose that it is a reasonable assumption that individuals contract the disease in accordance with a Poisson process, say $\{N(t), t \ge 0\}$, with parameter λ (probably unknown). That is, $N(t)$ is the number of people that have contracted the disease up to time t.

Let the distribution of the incubation time T be G. Assume that the incubation times are independent among infected people, and independent of the number of people infected. Let

- $N_1(t)$ denote the total number of people that have shown symptoms of the disease by time t, and
- $N_2(t)$ denote the total number of people that are infected, but have not shown symptoms of the disease by time t.

Consider a person that was infected at time s. The probability that this person shows symptoms of the disease by time t is equal to $P(T \le t - s) = G(t - s)$. In other words, N_1 counts an *event* (= infection) at time s with probability $G(t - s)$.

Then we know that that $N_1(t)$ and $N_2(t)$ are:

- Independent and Poisson distributed
- $N_1(t)$ has parameter

$$\lambda \int_0^t G(t-s)ds = \lambda \int_0^t G(s)ds.$$

- $N_2(t)$ has parameter

$$\lambda \int_0^t \overline{G}(t-s)ds = \lambda \int_0^t \overline{G}(s)ds.$$

Example 2.3(C): (p.70).

- Refer to the textbook for details.

E. Nonhomogeneous Poisson Process

E.1 Definition

Consider a counting process $\{N(t), t \geq 0\}$ having the following properties:

- $N(0) = 0$;
- Independent increments;
- $N(t + s) - N(t)$ is Poisson distributed with

$$\int_t^{t+s} \lambda(s)ds,$$

where $\lambda(t)$ is some nonnegative function of t.

Such a process is called ***nonstationary or nonhomogeneous Poisson process*** with *intensity function* $\lambda(t)$. For such a process, we define the *mean value function* $m(t)$ as follows:

$$m(t) = E(N(t)) = \int_0^t \lambda(s)ds$$

Using this notation, the random variable $N(t + s) - N(t)$ is Poisson distributed with parameter $m(t + s) - m(t)$, that is,

$$P(N(t+s) - N(t) = n) = e^{-(m(t+s)-m(t))} \frac{[m(t+s) - m(t)]^n}{n!}, \text{ for all } n \geq 0.$$

Note that if $\lambda(t) = \lambda$ for all i, we obtain the traditional, sometimes also called homogeneous Poisson process with rate or intensity λ.

An equivalent characterization is as follows (Ross uses this as the definition):

A counting process $\{N(t), t \geq 0\}$is *a nonhomogeneous Poisson process* with intensity function $\lambda(t)$ if and only if

1. $N(0) = 0$;
2. The process has independent increments;
3. $P(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$;
4. $P(N(t+h) - N(t) \geq 2) = o(h)$.

Fix t and define

$$P_n(s) = P(N(t + s) - N(t) = n).$$

For $n = 0$, we have

$$\begin{aligned} P_0(s + h) &= P(N(t + s + h) - N(t) = 0) \\ &= P(0 \text{ events in } (t, t+s), 0 \text{ events in } (t+s, t+s+h)) \\ &= P(0 \text{ events in } (t, t+s)) \cdot P(0 \text{ events in } (t+s, t+s+h)) \\ &= P_0(s) [1 - \lambda(s + t) h + o(h)] \end{aligned}$$

which implies that

$$P'_0(s) = - \lambda(s + t)P_0(s)$$

$$\Rightarrow \log P_0(s) = -\int_0^s \lambda(t+u)du \Rightarrow P_0(s) = e^{-(m(t+s)-m(t))}.$$

For general n, the proof follows in a similarly way.

E.2 Relationship between Homogenous and Nonhomogeneous Poisson Processes

A. Homogeneous $\Rightarrow$ Nonhomogeneous: Filtering or Sampling

Let $\lambda(t)$ be some nonnegative function such that $\lambda(t) \le \lambda$. Let $\{N'(t), t \ge 0\}$be a homogeneous Poisson process with rate λ. Consider the related process $\{N(t), t \ge 0\}$, obtained from $\{N'(t), t \ge 0\}$ by counting an event occurring at time t with probability $\lambda(t)/\lambda \in [0, 1]$.

We know that the random variables $N(t)$ are Poisson distributed with parameter $m(t)$, which follows from the discussion of "multiple event types" associated with $P(s) = \lambda(s)/\lambda$. In fact, $\{N(t), t \ge 0\}$ is a nonhomogeneous Poisson process with the intensity function $\lambda(t)$. This fact implies that a nonhomogeneous Poisson process with a bounded intensity function is a result of ***filtering*** of a homogeneous Poisson process.

How to prove this?

$$P(\text{ one counted event in } (t, t + h)) = P(\text{one event in } (t, t+s))\cdot(\lambda(t)/\lambda) + o(h)$$
$$= \lambda(t)h + o(h).$$

Example: **Output Process of *M*/*G*/**∞ (Example 2.4(B), p.81)

B. Nonhomogeneous $\Rightarrow$ Homogeneous: Rescaling of Time

Let $\{X(t): t \ge 0\}$ be a nonhomogeneous Poisson process with rate $\lambda(t)$. Define

$$\Lambda(t) = \int_0^t \lambda(u)du.$$

Now make a deterministic change in time scale and define a new process $\{N(s): s \ge 0\}$ as follows: $N(s) = X(t)$ when $s = \Lambda(t)$.

Note that $\Delta s = \lambda(t)\, \Delta t + o(\Delta t)$. Then we have

$$P(N(s+\Delta s) - N(s) = 1) = P(X(t+\Delta t) - X(t) = 1)$$
$$= \lambda(t)\Delta t + o(\Delta t) = \Delta s + o(\Delta s),$$

which implies that the new process $\{N(s): s \ge 0\}$ is a homogeneous Poisson process with rate 1.

E.3 Interarrival Times

Similar to the homogeneous Poisson process, it is easy to see that

$$P(X_1 > t) = P(N(t) = 0) = e^{-m(t)}$$

If $\lim_{t\to\infty} m(t) < \infty$, then $P(X = \infty) > 0$.

Now consider the random variable X_2. We can find its distribution by conditioning on X_1.

$$P(X_2 > t) = \int_0^\infty P(X_2 > t \mid X_1 = s)\lambda(s)e^{-m(s)}ds$$
$$= \int_0^\infty e^{-(m(t+s)-m(s))}\lambda(s)e^{-m(s)}ds = \int_0^\infty \lambda(s)e^{-m(t+s)}ds$$

Not surprisingly, X_1 and X_2 are not identically distributed. Furthermore, X_1 and X_2 are not independent as well since the conditional distribution of X_2 given $X_1 = s$ depends on s.

F. Compound Poisson Variables and Processes

F.1 Compound Poisson Random Variable

Let X_1, X_2, …. be a sequence of independent and identically distributed random variables having distribution function F, and suppose that this sequence is independent of N, a Poisson random variable with mean λ. The new random variable

$$W = \sum_{i=1}^{N} X_i$$

is said to be a ***compound Poisson*** random variable with Poisson parameter λ and component distribution F.

The m.g.f. of W is given by:

$$\phi_W(t) = E\left[e^{tW}\right] = \sum_{n=0}^{\infty} E\left[e^{tW} \middle| N = n\right] P(N = n) = \sum_{n=0}^{\infty} E\left[e^{t(X_1+\ldots+X_n)} \middle| N = n\right] e^{-\lambda} \frac{\lambda^n}{n!}$$

$$= \sum_{n=0}^{\infty} \left[E(e^{tX_1})\right]^n e^{-\lambda} \frac{\lambda^n}{n!} = \sum_{n=0}^{\infty} \left[\phi_X(t)\right]^n e^{-\lambda} \frac{\lambda^n}{n!} = \exp\{\lambda(\phi_X(t) - 1)\}$$

where $\phi_X(t)$ is the m.g.f. for X_i's.

Note that,

$$\phi_W\,'(t) = \lambda\phi_X\,'(t) e^{\lambda(\phi_X(t)-1)}$$

$$\phi_W\,''(t) = [\lambda\phi_X\,''(t) + (\lambda\phi_X\,'(t))^2] e^{\lambda(\phi_X(t)-1)}$$

which implies that

$$\phi_W\,'(0) = \lambda E(X), \quad \phi_W\,''(0) = \lambda E(X^2) + (\lambda E(X))^2$$

Therefore,

$$E(W) = \lambda\, E(X);\ Var(W) = \lambda\, E(X^2).$$

<u>Example (ATM)</u>: Consider the number of customers, say N, that draw money from an ATM during a given day. Let N be a Poisson random variable with mean λ. Let each person, say i, independently of all other persons, withdraw an amount of money, say X_i, is distributed according to a distribution function F. Let W be the total amount withdrawn during the day. Then W is a compound Poisson random variable.

F.2 Compound Poisson Process

Let $\{N(t), t \geq 0\}$ be a Poisson process, and let the i-th event cause a contribution of X_i. Let the sequence X_1, X_2,…be i.i.d., and independent of the Poisson process. Then the stochastic process $\{W(t), t \geq 0\}$, defined by:

$$W(t) = \sum_{i=1}^{N(t)} X_i$$

is called a ***compound Poisson process***.

Clearly, $W(t)$ is a compound Poisson random variable for all t. In terms of the above ATM example, if $\{N(t), t \geq 0\}$ counts the customers that have arrived up to time t, then the process $\{W(t), t \geq 0\}$ keeps track of the total amount withdrawn up to time t.

Example 2.5(A): Consider events that occur in accordance with a Poisson process with rate α: $\{N(t), t \geq 0\}$ An event occurring at time s, independent of the past of the process, contributes a random amount having distribution F_s. Let C_i be amount of contribution if event i occurs. Then the total amount of contributions up to time t, W, is given by:

$$W = \sum_{i=1}^{N(t)} C_i$$

First note that $N(t)$ has a Poisson distribution with parameter αt. If we condition on $N(t) = n$, the (unordered) event times are i.i.d. uniformly distributed in $[0, t]$. The distribution of the contribution of each of these events, say X, can be determined by conditioning on the event time, say T. The distribution of X is

$$F(x) = P(X \leq x) = \int_0^t P(X \leq x \mid T = s) f(s) ds = \int_0^t P(X \leq x \mid T = s) \frac{1}{t} ds = \frac{1}{t} \int_0^t F_s(x) ds.$$

Consequently, the total contribution W has the same distribution as:

$$W' = \sum_{i=1}^{N(t)} X_i$$

where $X_i \sim F$ as defined above and is therefore compound Poisson distributed with parameter αt and component distribution F.

G. Conditional Poisson Processes

G.1 Definition

Consider a Poisson process with *unknown* parameter λ. However, we do have some prior information in term of *prior distribution*. In other words, let Λ have some distribution G, and let a realization of this random variable be the rate of the Poisson process. This process, say $\{N(t), t \geq 0\}$, is called ***a conditional Poisson process***.

- Is it a Poisson process?
- Does the process have stationary increments?
- Does the process have independent increments?

Consider the distribution of its increments $N(t + s) - N(s)$:

$$P(N(t+s) - N(s) = n) = \int_0^\infty e^{-\lambda t} \frac{(\lambda t)^n}{n!} dG(\lambda).$$

which only depends on the value of t, i.e., the length of the interval $(s, s + t]$. This implies that the process has stationary increments. But the process $\{N(t), t \geq 0\}$ is not a Poisson process because the increments are not Poisson distributed. Furthermore, it does not have independent increments:

$$
\begin{aligned}
&P(N(t+s)-N(s)=n, N(s)=m) \\
&=\int_0^\infty P(N(t+s)-N(s)=n, N(s)=m \mid \Lambda=\lambda)dG(\lambda) \\
&=\int_0^\infty P(N(t+s)-N(s)=n \mid \Lambda=\lambda)\cdot P(N(s)=m \mid \Lambda=\lambda)dG(\lambda) \\
&\neq \int_0^\infty P(N(t+s)-N(s)=n \mid \Lambda=\lambda)dG(\lambda)\cdot\int_0^\infty P(N(s)=m \mid \Lambda=\lambda)dG(\lambda)
\end{aligned}
$$

Intuitively, this makes sense: Observing the events of the process should give us some information on the value of λ.

G.2 *Conditional Distribution of* Λ

We will study the conditional distribution of the random variable Λ, given the history of the Poisson process up to time t. he history can be summarized through $N(t) = n$, since Λ is the random rate, i.e., only the number of events per unit time is important. More formally, when we condition on $N(t) = n$, the unordered event times are uniform in [0, t] regardless of the value of λ, and are therefore irrelevant. Then,

$$
\begin{aligned}
P(\Lambda\le\lambda \mid N(t)=n) &= \frac{P(\Lambda\le\lambda, N(t)=n)}{P(N(t)=n)} \\
&=\int_0^\lambda \frac{P(N(t)=n \mid \Lambda=x)dG(x)}{P(N(t)=n)} = \frac{\int_o^\lambda (N(t)=n \mid \Lambda=x)dG(x)}{\int_o^\infty (N(t)=n \mid \Lambda=x)dG(x)} \\
&=\frac{\int_0^\lambda e^{-xt}\frac{(xt)^n}{n!}dG(x)}{\int_0^\infty e^{-xt}\frac{(xt)^n}{n!}dG(x)} = \frac{\int_0^\lambda e^{-xt}(xt)^n dG(x)}{\int_0^\infty e^{-xt}(xt)^n dG(x)}
\end{aligned}
$$

Hence the underlying density function is given by:

$$
\frac{e^{-\lambda t}(\lambda t)^n dG(\lambda)}{\int_0^\infty e^{-xt}(xt)^n dG(x)}
$$

<u>Example</u>: Gamma prior distribution

The Gamma distribution is often chosen as a prior distribution of Λ. Partly because of the fact that it has 2 parameters, and can approximate many practical nonnegative random variables; but also because of its attractive properties as a prior. In particular, let

$$
g(\lambda)=\alpha e^{-\lambda\alpha}\frac{(\lambda\alpha)^{m-1}}{(m-1)!} \text{ for } \lambda\ge 0.
$$

(Note: the Gamma distribution is also defined for non-integer values of the parameter m, and all results generalize to these cases as well.)

It can be shown that the density function of Λ given $N(t) = n$ at λ is given by

$$g(\lambda \mid N(t)=n) = (\alpha+t)e^{-\lambda(\alpha+t)} \frac{(\lambda(\alpha+t))^{n+m-1}}{(n+m-1)!}$$

which says that the conditional distribution of Λ given $N(t) = n$ is a Gamma distribution with parameter $n + m$ and $\alpha + t$.

Note that, a prior, we have

$$E(\Lambda) = \frac{m}{\alpha},$$

and after updating, that is, $N(t) = n$, we have

$$E(\Lambda) = \frac{m+n}{\alpha+t}.$$

We can rewrite this as,

$$E(\Lambda) = \frac{m+n}{\alpha+t} = \frac{m}{\alpha}\cdot\frac{\alpha}{\alpha+t} + \frac{n}{t}\cdot\frac{t}{\alpha+t},$$

which is the following weighted average of the prior expected value and the observed value. Note that the weight for the observed value increases to 1 as $t \to \infty$.

G.3 *Local Behavior of a Conditional Poisson Process*

How does the process behave locally, i.e., for a small amount of time after the current time t? That is, we are interested in the rate of the process at time t.

The rate of the process at time 0 is equal to

$$\lim_{h\downarrow 0}\frac{P(N(h)=1)}{h} = \lim_{h\downarrow 0}\frac{\int_0^\infty P(N(h)=1 \mid \Lambda=\lambda)dG(\lambda)}{h}$$

$$= \lim_{h\downarrow 0}\frac{\int_0^\infty (\lambda h + o(h))dG(\lambda)}{h} = \int_0^\infty \lambda dG(\lambda) = E(\Lambda).$$

Similarly, the rate of the process at time t is equal to

$$\lim_{h\downarrow 0}\frac{P(N(t+h)-N(t)=1 \mid N(t)=n)}{h}$$

$$= \lim_{h\downarrow 0}\frac{\int_0^\infty P(N(t+h)-N(t)=1 \mid \Lambda=\lambda)dG(\lambda \mid N(t)=n)}{h}$$

$$= \lim_{h\downarrow 0}\frac{\int_0^\infty (\lambda h + o(h))dG(\lambda \mid N(t)=n)}{h}$$

$$= \int_0^\infty \lambda dG(\lambda \mid N(t)=n) = E(\Lambda \mid N(t)=n).$$

Thus, a conditional Poisson process behaves locally as a Poisson process with parameter equal to its conditional expectation given the past of the process.

H. Additional Examples

Example 1: (**Coupon Collector Problem**) Consider the following situation:

- There is an (essentially) infinite population of coupons from which we sample;
- There are m distinct coupons;
- Coupon j is observed with probability p_j.

What is the expected number of times we need to sample a coupon from the population until we have observed ***all*** coupon types?

Pre-analysis: Note that we cannot follow the previous analysis (filtered process) since the coupon types are not equally likely to be observed. Another approach may be the following.

- Let N be the number of coupons needed to observe all types;
- Let N_j be the number of coupons needed to obtain one coupon of type j.
- Clearly, N_j has a geometric distribution with parameter p_j.
- Moreover, $N = \max_{\{j=1,..,m\}} N_j$

However, unfortunately the random variables N_j are not independent!

Let us try to study a related problem first. Assume that we collect coupons according to a Poisson process $\{N(t), t \geq 0\}$ with rate $\lambda = 1$. That is, we observe a coupon whenever an event of this process occurs. Thus, $N(t)$ now denotes the number of coupons observed by time t. In addition, let the counting process $\{N_j(t), t \geq 0\}$ count the number of coupons of type j observed by time t (j=1,…,m). These m processes are independent Poisson processes with rates $\lambda p_j = p_j$.

Now define the following random variables:

- Let X_j be the time at which we observe the first coupon of type j.
  - Due to the independence of the Poisson processes, these are independent exponential random variables with parameters p_j (j =1,…,m).
- Let X be time at which we have observed all coupon types: $X = \max_{\{1,\dots,m\}} X_i$.

The distribution of X can be derived as follows:

$$\Pr(X \leq t) = \Pr(\max_{j=1,\dots,m} X_j \leq t) = \Pr(X_j \leq t, \forall j = 1,\dots,m)$$

$$= \prod_{j=1}^{m} \Pr(X_j \leq t) = \prod_{j=1}^{m} (1 - e^{-p_j t}).$$

Then the expected value of X is given by:

$$E(X) = \int_0^\infty \Pr(X > t)dt = \int_0^\infty \left(1 - \prod\nolimits_{j=1}^{m} (1 - e^{-p_j t})\right)dt.$$

Now let us try fine $E(N)$. Let T_n denote the n-th interarrival time of the aggregate Poisson process $\{N(t), t \geq 0\}$. Then the relationship between X and N is given by

$$X = \sum_{n=1}^{N} T_n \,.$$

Now note that

- the random variables T_n are i.i.d. exponential distributed with parameter 1;
- the random variable N is independent of the T_n's.

Then, it follows that

$$E(X) = E(N)\cdot E(T_1) = E(N) = \int_0^\infty \left(1-\prod_{j=1}^m (1-e^{-p_j t})\right)dt.$$

Example 2: (**Estimating Software Reliability**) Consider a testing procedure for a new computer package, aiming to eliminate bugs from the package. Let us assume that we continue to run tests until some predetermined time *t*, and then remove all detected bugs from the program. What can we say about the error rate of the revised program?

To simplify matters, let us assume

- The package contains an unknown number *m* of bugs;
- Bug *i* will cause an error to occur in accordance with a Poisson process with unknown rate λ_i;
- These Poisson processes are independent.

Then the aggregate Poisson process counting the total number of errors then has rate $\sum_{j=1}^m \lambda_j$, which is the initial error rate of the package.

After running the testing procedure until time *t* and removing the detected bugs, the error rate of the revised package will be

$$\Lambda(t) = \sum_{i=1}^m \lambda_i I_{A_i(t)}$$

where $A_i(t)$ stands for the event that "bug *i* has not caused an error by time *t*". It is clear that $\Lambda(t)$ is a random variable and let us estimate this rate by computing its expected value.

Since $P(A_i(t)) = \Pr(N_i(t)=0) = e^{-\lambda_i t}$, it follows that

$$E(\Lambda(t)) = \sum_{i=1}^m \lambda_i e^{-\lambda_i t}$$

But the trouble is we have no information on the values of *m* and λ_i!

Now consider also the random variables $M_j(t)$, which denote the number of bugs that were responsible for *j* errors during [0, *t*]. In particular,

$$E(M_1(t)) = \sum_{i=1}^m \Pr(A_i^1(t)) = \sum_{i=1}^m \lambda_j t e^{-\lambda_i t} = tE(\Lambda(t)),$$

where $A_i^1(t)$ represents the event that "bug *i* has caused exactly 1 error by time *t*". This implies that

$$E(\Lambda(t)) = \frac{1}{t}E(M_1(t)),$$

i.e., we can use the observed value of $M(t)$ to estimate the error rate of the revised package.

Example 3: (**Elevators Problem**) Consider an elevator that starts in the basement of a building and travels upward. Let N_i denote the number of people that get in the elevator at floor *i*. Assume that the N_i's are independent Poisson distributed random variables with parameters λ_i.

Each person entering the elevator at floor i will, independent of everything else, get off the elevator at floor j with probability P_{ij} ($j > i$). Clearly,

$$\sum_{j:j>i} P_{ij} = 1, \text{ for all } i$$

Let O_j denote the number of people getting off the elevator at floor j. For simplicity, assume that the elevator has infinite capacity. Questions are:

- What is the expected value of O_j?
- What is the distribution of O_j?
- What is the joint distribution of O_j and O_k?

Let $\{N_i(t), t \geq 0\}$ be independent Poisson processes with parameters λ_i. Note that $N_i(1) \sim N_i$. Note that t here is an artificial parameter, and does not represent time! Now filter each of the Poisson processes as follows:

- An event is classified as being of Type $i.j$ with probability P_{ij} ($j > i$).

Then the ***filtered*** processes $\{N_{ij}(t), t \geq 0\}$ are *independent* Poisson processes with parameters λP_{ij}. Finally, let us *aggregate* these processes over i to obtain the following *independent* Poisson processes:

$$\left\{ O_j(t) \equiv \sum_{i:i<j} N_{ij}(t), t \geq 0 \right\}$$

with parameter $\sum_{i:i<j} \lambda_i P_{ij}$.

Now realize that $O_j(1) \sim O_j$ for all j. This implies that

$$O_j \sim Poisson\left(\sum_{i:i<j} \lambda_i P_{ij} \right).$$

In addition, since the Poisson processes are independent, O_j and O_k are independent for $j \neq k$.

Example 4: (**Infectious Disease Spreading – More Analysis**)

- Individuals contract the disease in accordance with a Poisson process, say $\{N(t), t \geq 0\}$, with ***unknown*** parameter λ. That is, $N(t)$ is the number of people that have contracted the disease up to time t.
- Let the (known) distribution of the incubation time T be G. Assume that the incubation times are independent among infected people, and independent of the number of people infected.
- $N_1(t)$ denote the total number of people that have shown symptoms of the disease by time t.
- $N_2(t)$ denote the total number of people that are infected, but have not shown symptoms of the disease by time t.

We have shown that $N_1(t)$ and $N_2(t)$ are independent and Poisson distributed, having the parameter

$$\lambda \int_0^t G(t-s)ds = \lambda \int_0^t G(s)ds, \text{ and } \lambda \int_0^t \bar{G}(t-s)ds = \lambda \int_0^t \bar{G}(s)ds$$

respectively.

Note that, as $t \to \infty$, the expected unobserved cases converges:

$$\lim_{t\to\infty} E(N_2(t)) = \lim_{t\to\infty} \lambda \int_0^t \overline{G}(s)ds = \lambda \int_0^\infty \overline{G}(s)ds = \lambda E(T)$$

Similarly,

$$\lim_{t\to\infty} \frac{E(N_1(t))}{\lambda t} = \lim_{t\to\infty} \frac{1}{t}\int_0^t G(s)ds = \lim_{t\to\infty} \frac{1}{t}\int_0^t (1-\overline{G}(s))ds = 1 - \lim_{t\to\infty} \frac{1}{t}\int_0^t \overline{G}(s)ds = 1.$$

Let's return to the original question: Can we say something about the total number of people that are infected with the disease? Recall

$$E(N_2(t)) = \lambda \int_0^t \overline{G}(s)ds.$$

However, we do not know the value of λ. Note that $E(N_1(t)) = \lambda \int_0^t G(s)ds$. If we observe $N_1(t) = n_1$, we can estimate λ as follows:

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

which yields

$$\hat{N}_2(t) = n_1 \times \int_0^t \overline{G}(s)ds \Big/ \int_0^t G(s)ds.$$

Now let us take a closer look at the (Poisson distributed) random variables $N_1(t)$. In particular, consider the counting process $\{N_1(t),\ t \geq 0\}$. That is, this process counts the number of people that have shown symptoms of the disease by time t.

- Is $\{N_1(t), t \geq 0\}$ a filtering of the Poisson process $\{N(t), t \geq 0\}$?
- Is $\{N_1(t), t \geq 0\}$ a Poisson process?
- If so, why, and what is its parameter?
- If not,
  - Does it have independent increments?
  - Does it have stationary increments?

Recall that $E(N_1(t)) = \lambda \int_0^t G(s)ds$. So $\{N_1(t), t \geq 0\}$ cannot be a Poisson process.

Before we answer the remaining two questions, let's first study it's relationship with $\{N(t), t \geq 0\}$ a bit more closely. Recall that we determined $E(N_1(t))$ by filtering the *infections*, counted by $N(t)$. More precisely, we filtered out the infections that *will not* show symptoms by time t. But the process $\{N_1(t), t \geq 0\}$ is *not* filtering of $\{N(t), t \geq 0\}$ because

- the events in the former are the start of symptoms;
- the events in the later are the infections.

More formally speaking, $\{N_1(t), t \geq 0\}$ is *not* a filtering of $\{N(t), t \geq 0\}$ since the filtering probabilities $P(s) = G(t - s)$ depends on t.

Does $\{N_1(t): t \geq 0\}$ have *independent increments*? That is, if $0 \leq t_0 < t_1 < \ldots < t_n$, are the random variables $N_1(t_1) - N_1(t_0), \ldots, N_1(t_n) - N_1(t_{n-1})$ independent?

In the original Poisson process $\{N(t), t \geq 0\}$, declare an event to be of type k if symptoms of the infection occur in $(t_{k-1}, t_k]$. With this, $N_1(t_k) - N_1(t_{k-1})$ is precisely the total number of events of type k of the original Poisson process. These random variable are *independent* (and Poisson distributed) using the filtering proposition. Thus, $\{N_1(t): t \geq 0\}$ have independent increments.

Does $\{N_1(t): t \geq 0\}$ have *stationary increments*? Consider the number of events taking place in the interval $(t, t + s]$, i.e., $N_1(t + s) - N_1(t)$. In this case, let events of $\{N(t), t \geq 0\}$ be type 1 events if symptoms occur in $(t, t + s]$. This implies that $N_1(t + s) - N_1(t)$ is Poisson distributed with parameter

$$\lambda \int_0^{t+s} \pi(y)dy$$

where $\pi(y)$ denotes the probability that an event taking place at time y is of type 1. Note that

$$\pi(y) = \begin{cases} G(t+s-y) - G(t-y) & \text{if } y \leq t \\ G(t+s-y) & \text{if } y > t \end{cases}$$

Thus the parameter of the Poisson distribution is

$$\begin{aligned} \lambda \int_0^{t+s} \pi(y)dy &= \lambda \left(\int_0^t (G(t+s-y) - G(t-y))dy + \int_t^{t+s} G(t+s-y)dy \right) \\ &= \lambda \left(\int_0^{t+s} G(t+s-y)dy - \int_0^t G(t-y)dy \right) = \lambda \left(\int_0^{t+s} G(y)dy - \int_0^t G(y)dy \right) \\ &= \lambda \int_t^{t+s} G(y)dy \end{aligned}$$

which depends on both t and s. Therefore, the increments are not stationary.

Summarizing the above discussions, the process $\{N_1(t): t \geq 0\}$ has the following properties:

- It is a counting process.
- It has independent increments.
- $N_1(t + s) - N_1(t)$ is Poisson distributed with parameter $\lambda \int_t^{t+s} G(y)dy$.

That is, $\{N_1(t): t \geq 0\}$ is nonhomogeneous Poisson process with intensity function $\lambda G(t)$.

Example 5: (**Record Values**) Let $X_1, X_2, \ldots$ denote a sequence of i.i.d. nonnegative continuous random variables with distribution function F, density function f, and *support* $(0, \infty)$.

- The *support* of a distribution is the *smallest* set A having the following property:

$$\int_A dF(x) = 1$$

We say that a *record* occurs at time n if $X_n > X_i$ for $i = 1, \ldots, n-1$. If a record occurs at time n, we say that X_n is a *record value*. We would like to address the following questions:

1. How many points do you need to sample in order to obtain a value exceeding t?

2. How many *record values* do you observe until observing a value exceeding t?

Solution: The number of points needed to obtain a value exceeding t has a geometric distribution with success probability $\bar{F}(t)$ and expected value $1/\bar{F}(t)$.

For part 2, Let $N(t)$denote the number of record values that have a value at most equal to t. Then $\{N(t), t \geq 0\}$is a counting process, where an event occurs if t is a record value. Note that t is not time in this case!

Consider the probability that at least one event takes place between t and $t + h$. This event takes place if the first value X_i that is larger than t is also smaller than $t + h$. Now realize that the random variables X_i are independent.

If F is the exponential distribution with rate λ, the desired probability is equal to $\approx \lambda h$ (due to the memoryless property). In general, we would like to know the probability that a sample point has value in $(t, t + h]$ given that it is a record value exceeding t (i.e., there is an "arrival" in the next h time units given that there was no event in the first t time units). This probability is thus approximately equal to $\lambda(t)h$, where

$$\lambda(t) = \frac{f(t)}{\bar{F}(t)}$$

is the failure rate of the distribution F. Note that the continuity of the distribution precludes the possibility of 2 simultaneous events.

If the increments are independent (why is this the case?), we can conclude that $\{N(t), t \geq 0\}$is a nonhomogeneous Poisson process with intensity function $\lambda(t)$. Now consider the mean value function of this process:

$$m(t) = \int_0^t \lambda(s)ds = \int_0^t \frac{f(s)}{\bar{F}(s)}ds = \int_1^{\bar{F}(t)} \left(-\frac{1}{z}\right)dz \quad (\text{with } z = \bar{F}(s))$$

$$= \int_{\bar{F}(t)}^{t} \frac{dz}{z} = \ln(z)\Big|_{\bar{F}(t)}^{1} = \ln\left(\frac{1}{\bar{F}(t)}\right).$$

In other words, the expected number of record values that are at most equal to t is equal to $\ln\left(1/\bar{F}(t)\right)$ and the expected number of record values that is needed to observe a value greater than t is equal to 1 + $\ln\left(1/\bar{F}(t)\right)$.

Remark:

- Note that the expected number of observed values is *exponentially larger* than the expected number of *record values* to reach the same goal.