Chapter I
The Poisson Process

1. Three Ways To Define The Poisson Process

A stochastic process $(N(t))_{t\geq 0}$ is said to be a *counting process* if $N(t)$ counts the total number of 'events' that have occurred up to time t. Hence, it must satisfy:

(i) $N(t) \geq 0$ for all $t \geq 0$.
(ii) $N(t)$ is integer-valued.
(iii) If $s < t$, then $N(s) \leq N(t)$.
(iv) For $s < t$, the increment $N((s,t]) \stackrel{\text{def}}{=} N(t) - N(s)$ equals the number of events that have occurred in the interval $(s,t]$.

A counting process is said to have *independent increments* if the numbers of events that occur in disjoint time intervals are independent, that is, the family $(N(I_k))_{1\leq k\leq n}$ consists of independent random variables whenever $I_1, ..., I_n$ forms a collection of pairwise disjoint intervals. In particular, $N(s)$ is independent of $N(s+t) - N(s)$ for all $s, t \geq 0$.

A counting process is said to have *stationary increments* if the distribution of the number of events that occur in any interval of time depends only on the length of the time interval. In other words, the process has stationary increments if the number of events in the interval $(s, s+t]$, i.e. $N((s, s+t])$ has the same distribution as $N((0,t])$ for all $s, t \geq 0$.

One of the most important types of counting processes is the Poisson process, which can be defined in various ways.

Definition 1.1. [The Axiomatic Way]. A counting process $(N(t))_{t\geq 0}$ is said to be a *Poisson process with rate (or intensity)* λ, $\lambda > 0$, if:

(PP1) $N(0) = 0$.
(PP2) The process has independent increments.
(PP3) The number of events in any time interval of length t is Poisson distributed with mean λt. That is, $N((s,t]) \stackrel{d}{=} Poi(\lambda t)$ for all $s, t \geq 0$:

$$\mathbb{P}(N((s,t]) = n) \;=\; e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad n \in \mathbb{N}_0.$$

If $\lambda = 1$, then $(N(t))_{t\geq 0}$ is also called *standard Poisson process.*

Note that condition (PP3) implies that $(N(t))_{t\geq 0}$ has stationary increments and also that

$$\mathbb{E}N(t) \;=\; \lambda t, \quad t \geq 0,$$

which explains why λ is called the rate of the process.

In order to determine if an arbitrary counting process is actually a Poisson process, the conditions (PP1–3) must be shown. Condition (PP1), which simply states that the counting of events begins at time $t = 0$, and condition (PP2) can usually be verified directly from our knowledge of the process. However, it is not at all clear how we could determine validity of condition (PP3), and for this reason an equivalent definition of a Poisson process would be useful.

A function $f : \mathbb{R} \to \mathbb{R}$ is said to be $o(h)$ (for $h \to 0$), if

$$\lim_{h\to 0} \frac{f(h)}{h} \;=\; 0.$$

Definition 1.2. [By Infinitesimal Description]. A counting process $(N(t))_{t\geq 0}$ is said to be a Poisson process with rate λ, $\lambda > 0$, if:

(PP1) $N(0) = 0$.
(PP4) The process has stationary and independent increments.
(PP5) $\mathbb{P}(N(h) = 1) = \lambda h + o(h)$.
(PP6) $\mathbb{P}(N(h) \geq 2) = o(h)$.

That the processes defined by 1.1 form a subclass of those defined by 1.2 is easily assessed, but a proof of the reverse inclusion requires some work which we postpone to the end of this section. However, the essence of the proof is disclosed by the following heuristic argument based upon the *Poisson limit theorem* which states that

$$\lim_{n\to\infty} B(n,\theta_n)(\{k\}) \;=\; Poi(\theta)(\{k\}), \quad k \in \mathbb{N}_0,$$

whenever $\theta, \theta_1, \theta_2, ...$ are positive numbers such that $n\theta_n \to \theta$, as $n \to \infty$ (☞ [1, Satz 29.4]).

Plainly, we must only argue that (PP1) and (PP4–6) ensure $N(t) \stackrel{d}{=} Poi(\lambda t)$ for all $t > 0$. To see this subdivide the interval $[0, t]$ into k equal parts where k is very large. Note that, by (PP6), the probability of having two or more events in any subinterval goes to 0 as $k \to \infty$. This follows from

$$\begin{aligned}
&\mathbb{P}(\text{2 or more events in any subinterval}) \\
&\quad \leq \sum_{i=1}^{k} \mathbb{P}(\text{2 or more events in the } i\text{th subinterval}) \\
&\quad = k\, o\left(\frac{t}{k}\right) \;=\; t\,\frac{o(t/k)}{t/k} \;\to\; 0
\end{aligned}$$

as $k \to \infty$. Hence, $N(t)$ will (with probability going to 1) just equal the number of subintervals in which an event occurs. However, by (PP4) this number will have a binomial distribution with parameters k and $p_k = \lambda t/k + o(t/k)$. By letting $k \to \infty$, we thus see that $N(t)$ will have a Poisson distribution with mean equal to

$$\lim_{k\to\infty} k\left[\lambda\frac{t}{k} + o\left(\frac{t}{k}\right)\right] \;=\; \lambda t + \lim_{k\to\infty}\left[t\frac{o(t/k)}{t/k}\right] \;=\; \lambda t.$$

The astute reader will have noticed the possibility that the previous two definitions may only be wishful thinking, in other words, that processes satisfying (PP1–6) do not exist. It is indeed the merit of our third constructive definition of a Poisson process that it settles the question of existence in an affirmative way.

Definition 1.3. [The Constructive Way]. A counting process $(N(t))_{t\ge 0}$ is said to be a *Poisson process with rate* λ, $\lambda > 0$, if

$$N(t) \;=\; \sum_{n\ge 1}\mathbf{1}_{(0,t]}(T_n), \quad t \ge 0, \tag{1.1}$$

for a sequence $(T_n)_{n\ge 1}$ having i.i.d. increments $Y_1, Y_2, ...$, say, with an $\mathrm{Exp}(\lambda)$-distribution. The T_n are called *jump or arrival epochs* and the Y_n *interarrival or sojourn times* associated with $(N(t))_{t\ge 0}$.

It is clear that any counting process $(N(t))_{t\ge 0}$ is completely determined by its associated sequence of jump epochs $(T_n)_{n\ge 1}$ via (1.1). Hence, the equivalence of Definitions 1.1 and 1.3 follows if one can show that in the case of i.i.d. $Y_1, Y_2, ...$ with $Y_1 \overset{d}{=} \mathrm{Exp}(\lambda)$, and thus $T_n \overset{d}{=} \Gamma(n,\lambda)$ for each $n \ge 1$, the conditions (PP1–3) are satisfied. While (PP1) holds trivially true, we note for (PP3) that

$$\begin{aligned}
\mathbb{P}(N(t) = n) \;&=\; \mathbb{P}(T_n \le t < T_{n+1}) \\
&=\; \mathbb{P}(T_n \le t < T_n + Y_{n+1}) \\
&=\; \int_0^t \mathbb{P}(Y_{n+1} > t - s)\,\Gamma(n,\lambda)(ds) \\
&=\; \int_0^t e^{-\lambda(t-s)}\,\frac{\lambda^n s^{n-1}}{(n-1)!}e^{-\lambda s}\;ds \\
&=\; \frac{\lambda^n}{(n-1)!}e^{-\lambda t}\int_0^t s^{n-1}\;ds \\
&=\; \frac{(\lambda t)^n}{n!}e^{-\lambda t}
\end{aligned}$$

for each $t > 0$ and $n \in \mathbb{N}$ ($\Rightarrow \mathbb{P}(N(t) = 0) = e^{-\lambda t}$). This shows $N(t) \overset{d}{=} Poi(\lambda t)$. Finally, it remains to argue that $(N(t))_{t\ge 0}$ has independent increments (condition (PP2)). The key to this is provided by the following lemma hinging on the lack of memory property of the exponential distribution. We state it without proof here.

Lemma 1.4. *If $(T_n)_{n\ge 1}$ has independent increments which are exponentially distributed* *with parameter* $\lambda > 0$*, then the sequence*

$$\boldsymbol{Z}(t) \stackrel{\text{def}}{=} (T_{N(t)+1} - t, T_{N(t)+2}, T_{N(t)+3}, ...)$$

is independent of $(N(t), T_1, ..., T_{N(t)})$ *and distributed as* $\boldsymbol{Z}(0) = (T_n)_{n\geq 1}$ *for every* $t \geq 0$*.*

Now, since $(N(s+t) - N(s))_{t\geq 0} = H(\boldsymbol{Z}(s))$ for some measurable function H and all $s \geq 0$, Lemma 1.4 implies the independence of $N(s)$ and $(N(s+t) - N(s))_{t\geq 0}$, in particular of $N(s)$ and $N(t_2) - N(t_1)$ for all $0 \leq s \leq t_1 \leq t_2$. The reader is invited to complete this argument to conclude (PP2).

Let us further note here that, given a counting process satisfying (PP1–3), the distribution of the first jump epoch T_1 follows immediately from

$$\mathbb{P}(T_1 > t) \;=\; \mathbb{P}(N(t) = 0) \;=\; e^{-\lambda t}$$

for all $t > 0$.

Finally, we will show now that Definition 1.2 does indeed imply Definition 1.1.

Proof of "1.2 ⇒ 1.1". Assuming (PP1) and (PP4–6), the task is to verify (PP3), i.e. $N(t) \stackrel{d}{=} Poi(\lambda t)$ for each $t > 0$. Put

$$P_n(t) \stackrel{\text{def}}{=} \mathbb{P}(N(t) = n)$$

and start by considering $P_0(t)$. We derive a differential equation for $P_0(t)$ in the following manner: For $t \geq 0$ and $h > 0$, we have

$$\begin{aligned}
P_0(t+h) &= \mathbb{P}(N(t+h) = 0) \\
&= \mathbb{P}(N(t) = 0, N(t+h) - N(t) = 0) \\
&= \mathbb{P}(N(t) = 0)\,\mathbb{P}(N(t+h) - N(t) = 0) \qquad (1.2)\\
&= P_0(t)P_0(h) \\
&= P_0(t)[1 - \lambda h + o(h)]
\end{aligned}$$

where the final three equations follow from (PP4) and the fact that (PP5) and (PP6) give $P_0(h) = \mathbb{P}(N(h) = 0) = 1 - \lambda h + o(h)$. Notice that the latter together with $P_0(t+h) = P_0(t)P_0(h)$ ensures $P_0(t) > 0$ for all $t > 0$. Replacing t with $t - h$ in (1.2), we also have

$$P_0(t) \;=\; P_0(t-h)[1 - \lambda h + o(h)]. \qquad (1.3)$$

It follows that $P_0(t)$ is continuous, as $P_0(t \pm h) \to P_0(t)$ for $h \to 0$. But (1.2) and (1.3) further yield

$$\frac{P_0(t+h) - P_0(t)}{h} \;=\; -\lambda P_0(t) + \frac{o(h)}{h}$$

as well as

$$\frac{P_0(t-h) - P_0(t)}{-h} \;=\; -\lambda P_0(t-h) + \frac{o(h)}{h}.$$

Again, by letting $h \to 0$ and using the continuity of $P_0(t)$, we infer

$$P_0'(t) \;=\; -\lambda P_0(t)$$

or

$$\frac{P_0'(t)}{P_0(t)} \;=\; -\lambda,$$

which implies, by integration,

$$\log P_0(t) \;=\; -\lambda t + c$$

or

$$P_0(t) \;=\; K e^{-\lambda t}.$$

Since $P_0(0) = \mathbb{P}(N(0) = 0) = 1$, we arrive at

$$P_0(t) \;=\; e^{-\lambda t}, \quad t \geq 0. \tag{1.4}$$

Turning to the case $n \geq 1$, we begin by noting that

$$\begin{aligned} P_n(t+h) &= \mathbb{P}(N(t+h) = n) \\ &= \mathbb{P}(N(t) = n, N(t+h) - N(t) = 0) \\ &\quad + \mathbb{P}(N(t) = n-1, N(t+h) - N(t) = 1) \\ &\quad + \mathbb{P}(N(t+h) = n, N(t+h) - N(t) \geq 2). \end{aligned}$$

By (PP6), the last term in the above is $o(h)$; hence, by using (PP4) and (PP5), we obtain

$$\begin{aligned} P_n(t+h) &= P_n(t)P_0(h) + P_{n-1}(t)P_1(h) + o(h) \\ &= (1 - \lambda h)P_n(t) + \lambda h P_{n-1}(t) + o(h). \end{aligned} \tag{1.5}$$

This and the same identity, but with t replaced by $t - h$, shows the continuity of $P_n(t)$ by an inductive argument. Rewriting (1.5) as

$$\frac{P_n(t+h) - P_n(t)}{h} \;=\; -\lambda P_n(t) + \lambda P_{n-1}(t) + \frac{o(h)}{h}$$

and further using the corresponding equation with $t - h$ in place of t, i.e.

$$\frac{P_n(t-h) - P_n(t)}{-h} \;=\; -\lambda P_n(t-h) + \lambda P_{n-1}(t-h) + \frac{o(h)}{h},$$

we obtain upon letting h tend to 0

$$P_n'(t) \;=\; -\lambda P_n(t) + \lambda P_{n-1}(t)$$

or, equivalently,

$$e^{\lambda t}[P_n'(t) + \lambda P_n(t)] \;=\; e^{\lambda t} \lambda P_{n-1}(t).$$

Hence,

$$\frac{d}{dt}[e^{\lambda t} P_n(t)] \;=\; e^{\lambda t} \lambda P_{n-1}(t). \tag{1.6}$$

Now use mathematical induction over n, the hypothesis being $P_{n-1}(t) = e^{-\lambda t}(\lambda t)^{n-1}/(n-1)!$, to infer from (1.6)

$$\frac{d}{dt}[e^{\lambda t}P_n(t)] \; = \; \frac{\lambda(\lambda t)^{n-1}}{(n-1)!}$$

implying that

$$e^{\lambda t}P_n(t) \; = \; \frac{(\lambda t)^n}{n!} + c.$$

Finally, since $P_n(0) = \mathbb{P}(N(0) = n) = 0$, we arrive at the desired conclusion

$$P_n(t) \; = \; e^{-\lambda t}\frac{(\lambda t)^n}{n!}, \quad t \geq 0.$$

This completes the proof of "1.2 $\Rightarrow$ 1.1". $\diamondsuit$

2. Conditional Distribution Of The Jump Epochs

Suppose we are told that exactly one event of a Poisson process has taken place by time t, and we are asked to determine the distribution of the time at which the event occured. Since a Poisson process possesses stationary and independent increments, it seems reasonable that each interval in [0,t] of equal length should have the same probability of containing the event. In other words, the time of the event should be uniformly distributed over $[0, t]$. This is easily checked since, for $s \leq t$,

$$\begin{aligned}
\mathbb{P}(T_1 \leq s | N(t) = 1) \; &= \; \frac{\mathbb{P}(T_1 \leq s, N(t) = 1)}{\mathbb{P}(N(t) = 1)} \\
&= \; \frac{\mathbb{P}(1 \text{ event in } (0, s], 0 \text{ events in } (s, t])}{\mathbb{P}(N(t) = 1)} \\
&= \; \frac{\mathbb{P}(1 \text{ event in } (0, s])\, \mathbb{P}(0 \text{ events in } (s, t])}{\mathbb{P}(N(t) = 1)} \\
&= \; \frac{\lambda s e^{-\lambda s} e^{-\lambda(t-s)}}{\lambda t e^{-\lambda t}} \\
&= \; \frac{s}{t}.
\end{aligned}$$

This result may be generalized, but before doing so we need to introduce the concept of order statistics.

Let $Y_1, ..., Y_n$ be n random variables. We say that $(Y_{(1)}, ..., Y_{(n)})$ is the *order statistic* corresponding to $(Y_1, ..., Y_n)$ if $Y_{(k)}$ is the kth smallest among $Y_1, ..., Y_n$, $k = 1, ..., n$. If the Y_i's are i.i.d. continuous random variables with probability density f, then the joint density of the order statistics is given by

$$f_{(\cdot)}(y_1, ..., y_n) \; = \; n! \prod_{i=1}^{n} f(y_i)\, \mathbf{1}_{\mathcal{S}}(y_1, ..., y_n), \tag{2.1}$$

where $\mathcal{S} \stackrel{\text{def}}{=} \{(s_1, ..., s_n) \in \mathbb{R}^n : s_1 < s_2 < ... < s_n\}$. The above follows because

(i) $(Y_{(1)}, ..., Y_{(n)})$ will equal $(y_1, ..., y_n) \in \mathcal{S}$ if $(Y_1, ..., Y_n)$ is equal to any of the $n!$ permutations of $(y_1, ..., y_n)$ and

(ii) the probability density of $(Y_1, ..., Y_n)$ at $(y_{i_1}, ..., y_{i_n})$ equals $f(y_{i_1})f(y_{i_2}) \cdot ... \cdot f(y_{i_n}) = \prod_{i=1}^n f(y_i)$ when $(i_1, ..., i_n)$ is a permutation of $(1, ..., n)$.

By treating densities as if they were probabilities, we then indeed obtain

$$\begin{aligned} \mathbb{P}(Y_{(1)} = y_1, ..., Y_{(n)} = y_n) &= \sum_{(i_1,...,i_n)} \mathbb{P}(Y_1 = y_{i_1}, ..., Y_n = y_{i_n}) \\ &= \sum_{(i_1,...,i_n)} f(y_{i_1}) \cdot ... \cdot f(y_{i_n}) \\ &= n! \prod_{i=1}^n f(y_i), \quad (y_1, ..., y_n) \in \mathcal{S}, \end{aligned}$$

where summation is over all permutations $(i_1, ..., i_n)$ of $(1, ..., n)$.

If the Y_i, $i = 1, ..., n$, are uniformly distributed over $(0, t)$, then it follows from the above that the joint density function of the order statistics is given by

$$f_{(\cdot)}(y_1, ..., y_n) = \frac{n!}{t^n} \mathbf{1}_{\mathcal{S}}(y_1, ..., y_n). \tag{2.2}$$

We are now ready for the following useful theorem.

Theorem 2.1. *Given that $N(t) = n$, the n jump epochs $T_1, ..., T_n$ have the same distribution as the order statistics corresponding to n independent random variables uniformly distributed on the interval $(0, t)$.*

Proof. We shall compute the conditional density function of $T_1, ..., T_n$ given that $N(t) = n$. So let $0 < t_1 < ... < t_n < t_{n+1} = t$ and let h_i be small enough so that $t_i + h_i < t_{i+1}$ for $i = 1, ..., n$. Now,

$$\begin{aligned} &\mathbb{P}(t_i < T_i \le t_i + h_i, i = 1, ..., n | N(t) = n) \\ &= \frac{\mathbb{P}(\text{exactly 1 event in } (t_i, t_i + h_i], i = 1, ..., n, \text{no events elsewhere in } (0, t])}{\mathbb{P}(N(t) = n)} \\ &= \frac{\lambda h_1 e^{-\lambda h_1} \cdot ... \cdot \lambda h_n e^{-\lambda h_n} e^{-\lambda(t-h_1-h_2-...-h_n)}}{e^{-\lambda t}(\lambda t)^n / n!} \\ &= \frac{n!}{t^n} h_1 \cdot h_2 \cdot ... \cdot h_n. \end{aligned}$$

Consequently,

$$\frac{\mathbb{P}(t_i < T_i \le t_i + h_i, i = 1, ..., n | N(t) = n)}{h_1 \cdot h_2 \cdot ... \cdot h_n} = \frac{n!}{t^n},$$

and by letting $h_i \to 0$, we obtain that the conditional density of $T_1, ..., T_n$ given that $N(t) = n$ is

$$f_{(\cdot)}(t_1, ..., t_n) = \frac{n!}{t^n}, \quad 0 < t_1 < ... < t_n,$$

which completes the proof. ◊

Before turning to an example, let us point out that the above result suggests the following efficient way of simulating a Poisson process on a time interval $[0,t]$:

(i) Generate a random number N having a Poisson distribution with mean λt.

(ii) If $N = n \geq 1$, then generate n random numbers $U_1, ..., U_n$ with a uniform distribution on $(0,1)$ and choose

$$(T_1, ..., T_n) \stackrel{\text{def}}{=} (tU_{(1)}, ..., tU_{(n)}) \tag{2.3}$$

as the arrival times in $(0,t)$.

Example 2.2. Suppose that travelers arrive at a train depot in accordance with a Poisson process with rate λ. If the train departs at time t, let us compute the expected sum of the waiting times of travelers arriving in $(0,t)$. That is, we want $\mathbb{E}[\sum_{i=1}^{N(t)}(t-T_i)]$ where T_i is the arrival time of the ith traveler. Conditioning on $N(t)$ yields

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{N(t)}(t-T_i)\middle|N(t)=n\right] &= \mathbb{E}\left[\sum_{i=1}^{n}(t-T_i)\middle|N(t)=n\right] \\
&= nt - \mathbb{E}\left[\sum_{i=1}^{n}T_i\middle|N(t)=n\right].
\end{aligned}$$

Now if we let $U_1, ..., U_n$ be independent random variables with a uniform distribution on $(0,1)$, then

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n}T_i\middle|N(t)=n\right] &= \mathbb{E}\left[\sum_{i=1}^{n}tU_{(i)}\right] && \text{(by Theorem 2.1, ☞ (2.3))} \\
&= t\,\mathbb{E}\left[\sum_{i=1}^{n}U_i\right] && \left(\text{since } \sum_{i=1}^{n}U_{(i)} = \sum_{i=1}^{n}U_i\right) \\
&= \frac{nt}{2}.
\end{aligned}$$

Hence,

$$\mathbb{E}\left[\sum_{i=1}^{N(t)}(t-T_i)\middle|N(t)=n\right] = nt - \frac{nt}{2} = \frac{nt}{2}$$

and

$$\mathbb{E}\left[\sum_{i=1}^{N(t)}(t-T_i)\right] = \frac{t}{2}\,\mathbb{E}N(t) = \frac{\lambda t^2}{2}.$$ ♠

Tagging. As an important application of Theorem 2.1 suppose that each event of a Poisson process with rate λ is classified (”tagged”) as being either a type I or type II event, and suppose that the probability of an event being classified as type I depends on the time at which it occurs. Specifically, suppose that if an event occurs at time s, then, independently of all else, it is classified as being a type I event with probability $P(s)$ and a type II event with probability $1-P(s)$. By using Theorem 2.1 we can prove the following proposition.

Proposition 2.3. *If $N_i(t)$ represents the number of type i events that occur by time t, $i=1,2$, then $N_1(t)$ and $N_2(t)$ are independent Poisson random variables having respective means λpt and $\lambda(1-p)t$, where*

$$p \overset{\text{def}}{=} \frac{1}{t}\int_0^t P(s)\ ds.$$

PROOF. As ususal, denote by $T_1, T_2, ...$ the successive jump epochs of $(N(t))_{t\geq 0}$ and let $I_1, I_2, ...$ be Bernoulli variables such that I_k equals 1 or 0 depending on whether the kth occurring event is classified as a type I or type II event. In accordance with the above description the I_k are conditionally independent given $(T_n)_{n\geq 1}$, and

$$\mathbb{P}(I_k=1|T_1=t_1,T_2=t_2,...) \ = \ \mathbb{P}(I_k=1|T_k=t_k) \ = \ P(t_k)$$

for $0<t_1<t_2<...$ Fix any $n\geq 1$ and let $(T_k^*,I_k^*)_{1\leq k\leq n}$ be a random shuffle of $(T_k,I_k)_{1\leq k\leq n}$. Then

$$\begin{aligned}
&\mathbb{P}(I_1^*=i_1,...,I_n^*=i_n|T_1^*=t_1,...,T_n^*=t_n)\\
&\qquad = \frac{1}{n!}\sum_\pi \mathbb{P}(I_{\pi(1)}=i_1,...,I_{\pi(n)}=i_n|T_{\pi(1)}=t_1,...,T_{\pi(n)}=t_n)\\
&\qquad = \frac{1}{n!}\sum_\pi \prod_{k=1}^n \mathbb{P}(I_{\pi(k)}=i_k|T_{\pi(k)}=t_k) \qquad (2.4)\\
&\qquad = \prod_{k=1}^n [P(t_k)^{i_k}(1-P(t_k))^{1-i_k}]
\end{aligned}$$

for any $i_1,...,i_n\in\{0,1\}$, where the summation is over all permutations π of $1,...n$. So the I_k^*, $k=1,...,n$, are conditionally independent given $T_1^*,...,T_n^*$, and

$$\mathbb{P}(I_k^*=1|T_1^*=t_1,...,T_n^*=t_n) \ = \ \mathbb{P}(I_k^*=1|T_k^*=t_k) \ = \ P(t_k)$$

for each $k=1,...,n$. We now prove that $I_1^*,...,I_n^*$ conditioned upon $N(t)=n$ are i.i.d. with $\mathbb{P}(I_k^*=1|N(t)=n)=p$. By Theorem 2.1, $T_1^*,...,T_n^*$ conditioned upon $N(t)=n$ are i.i.d. with a uniform distribution on $(0,t)$. By combining this with (2.4), we infer for $i_1,...,i_n\in\{0,1\}$

$$\begin{aligned}
&\mathbb{P}(I_1^*=i_1,...,I_n^*=i_n|N(t)=n)\\
&\qquad = \mathbb{P}(I_1=i_1,...,I_n=i_n|N(t)=n)\\
&\qquad = \int_{(0,t)^n}\mathbb{P}(I_1^*=i_1,...,I_n^*=i_n|T_1^*=t_1,...,T_n^*=t_n)\\
&\qquad\qquad\qquad \times\mathbb{P}(T_1^*\in dt_1,...,T_n^*\in dt_n|N(t)=n)\\
&\qquad = \frac{1}{t^n}\int_0^t\cdots\int_0^t\prod_{j=1}^n\left[P(t_j)^{i_j}(1-P(t_j))^{1-i_j}\right]\,dt_n\ldots dt_1
\end{aligned}$$

$$= \left[\frac{1}{t}\int_0^t P(s)\ ds\right]^s \left[1 - \frac{1}{t}\int_0^t P(s)\ ds\right]^{n-s} \qquad (s \stackrel{\text{def}}{=} i_1 + ... + i_n)$$
$$= p^s(1-p)^{n-s}$$

which proves the above claim. But this result in combination with the observation that $N_1(t) = \sum_{k=1}^{N(t)} I_k = \sum_{k=1}^{N(t)} I_k^*$ implies

$$\begin{aligned}
\mathbb{P}(N_1(t) = m, &N_2(t) = n | N(t) = m+n) \\
&= \mathbb{P}(N_1(t) = m | N(t) = m+n) \\
&= \mathbb{P}\left(\sum_{k=1}^{N(t)} I_k^* = m \middle| N(t) = m+n\right) \\
&= \binom{m+n}{m} p^m (1-p)^n
\end{aligned}$$

for all $m, n \in \mathbb{N}_0$, and so

$$\begin{aligned}
\mathbb{P}(N_1(t) = m, N_2(t) = n) &= \binom{m+n}{m} p^m (1-p)^n e^{-\lambda t} \frac{(\lambda t)^{m+n}}{(m+n)!} \\
&= \left[e^{-\lambda p t} \frac{(\lambda p t)^m}{m!}\right] \left[e^{-\lambda(1-p)t} \frac{(\lambda(1-p)t)^m}{m!}\right]
\end{aligned}$$

which completes the proof. ♢

The importance of the above proposition, which we will extend to the counting processes $(N_1(t))_{t\geq 0}$ and $(N_2(t))_{t\geq 0}$ in Theorem 3.4, is illustrated by the following example.

Example 2.4. ***The Infinite Server Poisson Queue.*** Suppose that customers arrive at a service station in accordance with a Poisson process with rate λ. Upon arrival the customer is immediately served by one of an infinite number of possible servers, and the service times are assumed to be independent with a common distribution G.

To compute the joint distribution of the number of customers that have completed their service and the number of customers that are in service (queue length) at t, call an entering customer a type I customer if it completes its service by time t and a type II customer if it does not complete service by time t. Now, if the customer enters at time s, $s \leq t$, then it will be a type I customer if its service time is less than $t - s$, and since the service time distribution is G, the probability of this will be $G(t-s)$. Hence,

$$P(s) = G(t-s), \quad s \leq t,$$

and from Proposition 2.3 we obtain that the distribution of $N_1(t)$ – the number of customers that have completed service by time t – is Poisson with mean

$$\mathbb{E}N_1(t) = \lambda \int_0^t G(t-s)\ ds = \lambda \int_0^t G(y)\ dy.$$

Similarly, $N_2(t)$, the number of customers being served at time t, is Poisson distributed with mean

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

Further $N_1(t)$ and $N_2(t)$ are independent. ♠

Example 2.5. Suppose that a device is subject to shocks that occur in accordance with a Poisson process having rate λ. The ith shock gives rise to a damage D_i. The D_i, $i \geq 1$, are assumed to be i.i.d. and also to be independent of $(N(t))_{t\geq 0}$, where $N(t)$ is the number of shock in $[0, t]$. The damage due to a shock is assumed to decrease exponentially in time. That is, if a shock causes an initial damage D, then a time t later its damage is $De^{-\alpha t}$.

If we suppose that the damages are additive, then $D(t)$, the damage at t, can be expressed as

$$D(t) \;=\; \sum_{i=1}^{N(t)} D_i e^{-\alpha(t-T_i)},$$

where T_i represents the arrival time of the ith shock. We can determine $\mathbb{E}D(t)$ as follows:

$$\begin{aligned}
\mathbb{E}[D(t)|N(t)=n] \;&=\; \mathbb{E}\left[\sum_{i=1}^{N(t)} D_i e^{-\alpha(t-T_i)}\middle| N(t)=n\right] \\
&=\; \mathbb{E}\left[\sum_{i=1}^{n} D_i e^{-\alpha(t-T_i)}\middle| N(t)=n\right] \\
&=\; \sum_{i=1}^{n} \mathbb{E}[D_i e^{-\alpha(t-T_i)}|N(t)=n] \\
&=\; \mathbb{E}D \sum_{i=1}^{n} \mathbb{E}\big[e^{-\alpha(t-T_i)}|N(t)=n\big] \\
&=\; \mathbb{E}D\, \mathbb{E}\left[\sum_{i=1}^{n} e^{-\alpha(t-T_i)}\middle| N(t)=n\right] \\
&=\; \mathbb{E}D\, e^{-\alpha t}\, \mathbb{E}\left[\sum_{i=1}^{n} e^{\alpha T_i}\middle| N(t)=n\right].
\end{aligned}$$

Now, letting $U_1, ..., U_n$ be once again be i.i.d. uniform variables on $(0, 1)$, we obtain by another appeal to Theorem 2.1

$$\begin{aligned}
\mathbb{E}\left[\sum_{i=1}^{n} e^{\alpha T_i}\middle| N(t)=n\right] \;&=\; \mathbb{E}\left[\sum_{i=1}^{n} e^{\alpha t U_{(i)}}\right] \\
&=\; \mathbb{E}\left[\sum_{i=1}^{n} e^{\alpha t U_i}\right] \\
&=\; n\int_0^1 e^{\alpha t x}\; dx \;=\; \frac{n}{\alpha t}(e^{\alpha t}-1).
\end{aligned}$$

Hence,

$$\mathbb{E}[D(t)|N(t)] \;=\; \frac{N(t)}{\alpha t}(1-e^{-\alpha t})\,\mathbb{E}D$$

and, taking expectations,

$$\mathbb{E}D(t) \;=\; \frac{\lambda\,\mathbb{E}D}{\alpha}(1-e^{-\alpha t}). \qquad \spadesuit$$

Remark. Another approach to obtaining $\mathbb{E}D(t)$ is to break up the interval $(0,t]$ into nonoverlapping intervals of length h and then add the contribution at time t of shocks originating in these intervals. More specifically, let h be given and define X_i as the sum of the damages at time t of all shocks arriving in the interval $I_i \stackrel{\text{def}}{=} (ih,(i+1)h]$, $i=0,1,...,[t/h]$, where $[a]$ denotes the largest integer less than or equal to a. Then we have the representation

$$D(t) \;=\; \sum_{i=0}^{[t/h]} X_i,$$

and so

$$\mathbb{E}D(t) \;=\; \sum_{i=0}^{[t/h]} \mathbb{E}X_i.$$

To compute $\mathbb{E}X_i$ condition on whether or not a shock arrives in the interval I_i. This yields (recalling Definition 1.2)

$$\mathbb{E}D(t) \;=\; \sum_{i=0}^{[t/h]} (\lambda h\,\mathbb{E}[De^{-\alpha(t-L_i)}] + o(h)],$$

where L_i is the arrival time of the shock in the interval I_i. Hence,

$$\mathbb{E}D(t) \;=\; \lambda\,\mathbb{E}D\,\mathbb{E}\left[\sum_{i=0}^{[t/h]} he^{-\alpha(t-L_i)}\right] \;+\; \left[\frac{t}{h}\right]o(h). \tag{2.5}$$

But since $L_i \in I_i$, it follows upon letting $h \to 0$ that

$$\sum_{i=0}^{[t/h]} he^{-\alpha(t-L_i)} \;\to\; \int_0^t e^{-\alpha(t-y)}\;dy \;=\; \frac{1-e^{-\alpha t}}{\alpha}$$

and thus from (2.5) upon letting $h \to 0$

$$\mathbb{E}D(t) \;=\; \frac{\lambda\,\mathbb{E}D}{\alpha}(1-e^{-\alpha t}).$$

It is worth noting that the above is a more rigorous version of the following argument: Since the shock occurs in the interval $(y, y+dy]$ with probability λdy and since its damage at time t will equal $e^{-\alpha(t-y)}$ times its initial damage, it follows that the expected damage at t from shocks originating in $(y, y+dy]$ is

$$\lambda\,dy\,\mathbb{E}D\,e^{-\alpha(t-y)},$$

and so

$$\mathbb{E}D(t) \;=\; \lambda\,\mathbb{E}D \int_0^t e^{-\alpha(t-y)}\,dy \;=\; \frac{\lambda\,\mathbb{E}D}{\alpha}\,(1-e^{-\alpha t}). \qquad \spadesuit$$

2.1. The M/G/1 Busy Period

Consider the queueing system, known as $M/G/1$, in which customers arrive in accordance with a Poisson process with rate λ. Upon arrival they either enter service if the server is free or else they join the queue. The successive service times are independent and identically distributed according to G, and are also independent of the arrival process. When an arrival finds the server free, we say that a busy period begins. It ends when there are no longer any customers in the system. We would like to compute the distribution of the length of a busy period.

Suppose that a busy period has just begun at some time, which we shall designate as time 0. Let T_k denote the time until k additional customers have arrived. Thus T_k has a Gamma distribution with parameters k, λ [$T_k \stackrel{d}{=} \Gamma(k,\lambda)$]. Also let $X_1, X_2, ...$ denote the sequence of service times and put $S_k \stackrel{\text{def}}{=} X_1 + ... + X_k$. Now the busy period will last a time t and will consist of n services if, and only if,

(i) $T_k \le S_k$, $k = 1, ..., n-1$.

(ii) $S_n = t$.

(iii) There are $n-1$ arrivals in $(0,t)$.

Equation (i) is necessary for, if $T_k > S_k$, then the kth arrival after the initial customer will find the system empty of customers and thus the busy period would have ended prior to $k+1$ (and thus prior to n) services. The reasoning behind (ii) and (iii) is straightforward and left to the reader.

Hence, reasoning heuristically (by treating densities as if they were probabilities) we see from the above that

$$\begin{aligned}
&\mathbb{P}(\text{busy period is of length } t \text{ and consists of } n \text{ services})\\
&\quad = \mathbb{P}(S_n = t, n-1 \text{ arrivals in } (0,t), T_k \le S_k \text{ for } k = 1, ..., n-1)\\
&\quad = \mathbb{P}(T_k \le S_k \text{ for } k = 1, ..., n-1 | n-1 \text{ arrivals in } (0,t), S_n = t)\\
&\qquad\qquad \times \mathbb{P}(n-1 \text{ arrivals in } (0,t), S_n = t).
\end{aligned} \tag{2.6}$$

Now the arrival process is independent of the service times and thus

$$\mathbb{P}(n-1 \text{ arrivals in } (0,t), S_n = t) \;=\; e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}\,G_n(dt), \tag{2.7}$$

where G_n is the n-fold convolution of G with itself (the distribution of S_n). In addition, we have from Theorem 2.1 that, given $n-1$ arrivals in $(0,t)$, the ordered arrival times are distributed as the ordered values of a set of $n-1$ independent uniform$(0,t)$ random variables. Hence, using this fact along with (2.6) and (2.7) yields

$$\begin{aligned} &\mathbb{P}(\text{busy period is of length } t \text{ and consists of } n \text{ services}) \\ &\quad = e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!}\,G_n(dt)\,\mathbb{P}(T_k^* \le S_k,\ k=1,...,n-1|S_n=t), \end{aligned} \tag{2.8}$$

where $T_1^*, ..., T_{n-1}^*$ are independent of $X_1, ..., X_n$ and represent the ordered values of a set of $n-1$ uniform$(0,t)$ random variables.

To compute the remaining probability in (2.8) we need some lemmata. Given i.i.d. uniform$(0,1)$ random variables $U_1, ..., U_n$, we shall denote by $(U_{k:1}, ..., U_{k:k})$ the order statistic based on the first k of these variables $(1 \le k \le n)$. Lemma 2.6 is elementary and its proof is left as an exercise.

Lemma 2.6. *Let $S_1, ..., S_n$ be the partial sums of n i.i.d. nonnegative random variables $X_1, ..., X_n$. Then*

$$\mathbb{E}[S_k|S_n=s] \;=\; \frac{k}{n}s, \quad k=1,...,n.$$

Lemma 2.7. *Let $S_1, ..., S_n$ be as in Lemma 2.6 and $U_1, ..., U_n$ be the i.i.d. uniform$(0,1)$ random variables that are also independent of $S_1, ..., S_n$. Then*

$$\mathbb{P}(S_k \le tU_{n:k} \textit{ for } k=1,...,n|S_n=s) \;=\; \begin{cases} 1-\dfrac{s}{t}, & \textit{if } 0<s\le t, \\ 0, & \textit{otherwise} \end{cases} \tag{2.9}$$

for each $t>0$.

Proof. The proof is by induction on n. For $n=1$ we must compute $\mathbb{P}(S_1 \le tU_1|S_1=s)$. But

$$\mathbb{P}(S_1 \le tU_1|S_1=s) \;=\; \mathbb{P}\left(U_1 \ge \frac{s}{t}\right) \;=\; 1-\frac{s}{t}.$$

So assume the lemma be true if n is replaced by $n-1$ and consider the n case. Since the result is obvious for $s>t$, suppose that $s \le t$. To make use of the induction hypothesis we will compute the left-hand side of (2.9) by conditioning on the values of S_{n-1} and $U_{n:n}$ and then using the quite intuitive fact that

$$\mathbb{P}((U_{n:1}, ..., U_{n:n-1}) \in \cdot|U_{n:n}=u) \;=\; \mathbb{P}((uU_{n-1:1}, ..., uU_{n-1:n-1}) \in \cdot).$$

Doing so, we have for $y \le s$

$$\begin{aligned} &\mathbb{P}(S_k \le tU_{n:k},\ k=1,...,n|S_{n-1}=y, U_{n:n}=u, S_n=s) \\ &\quad = \mathbb{P}(S_k \le tU_{n:k},\ k=1,...,n|S_{n-1}=y, U_{n:n}=u, X_n=s-y) \\ &\quad = \begin{cases} \mathbb{P}(S_k \le tuU_{n-1:k},\ k=1,...,n-1|S_{n-1}=y), & \text{if } s \le tu \\ 0, & \text{otherwise.} \end{cases} \end{aligned}$$

$$= \begin{cases} 1 - \dfrac{y}{tu}, & \text{if } s \le tu, \\ 0, & \text{otherwise.} \end{cases} \qquad \text{(induction hypothesis)}$$

Hence, for $s \le tu$,

$$\mathbb{P}(S_k \le tU_{n:k},\ k = 1, ..., n | S_{n-1}, U_{n:n} = u, S_n = s) \ = \ 1 - \frac{S_{n-1}}{tu}$$

and thus, for $s \le tu$,

$$\begin{aligned} &\mathbb{P}(S_k \le tU_{n:k},\ k = 1, ..., n | U_{n:n} = u, S_n = s) \\ &= \ \mathbb{E}\left[1 - \frac{S_{n-1}}{tu} \middle| U_{n:n} = u, S_n = s\right] \\ &= \ \mathbb{E}\left[1 - \frac{S_{n-1}}{tu} \middle| S_n = s\right] \\ &= \ 1 - \frac{1}{tu}\,\mathbb{E}[S_{n-1}|S_n = s] \\ &= \ 1 - \frac{1}{tu}\frac{n-1}{n}s, \end{aligned}$$

where we have made use of Lemma 2.6 in the above. Taking expectations once more yields

$$\begin{aligned} &\mathbb{P}(S_k \le tU_{n:k},\ k = 1, ..., n | S_n = s) \\ &= \ \int_{(s/t,1)} \left(1 - \frac{s}{tu}\frac{n-1}{n}\right) \mathbb{P}(U_{n:n} \in du). \end{aligned} \tag{2.10}$$

Now the distribution function of $U_{n:n} = \max_{1 \le i \le n} U_i$ is given by

$$\mathbb{P}(U_{n:n} \le u) \ = \ \mathbb{P}(U_i \le u,\ i = 1, ..., n) \ = \ u^n, \quad 0 < u < 1,$$

and its density is therefore

$$f_n(u) \ = \ nu^{n-1}\mathbf{1}_{(0,1)}(u).$$

In (2.10), this leads to

$$\begin{aligned} &\int_{(s/t,1)} \left(1 - \frac{s}{tu}\frac{n-1}{n}\right) \mathbb{P}(U_{n:n} \in du) \\ &= \ \int_{s/t}^{1} \left(1 - \frac{s}{tu}\frac{n-1}{n}\right) nu^{n-1}\ du \\ &= \ 1 - \left(\frac{s}{t}\right)^n - \frac{s}{t}\left[1 - \left(\frac{s}{t}\right)^{n-1}\right] \ = \ 1 - \frac{s}{t} \end{aligned}$$

and the proof is complete. ◊

Lemma 2.8. *Under the same conditions as in Lemma 2.7 it holds true that*

$$\mathbb{P}(S_k \le tU_{n-1:k} \text{ for } k = 1, ..., n-1 | S_n = t) \ = \ \frac{1}{n}$$

for all $t > 0$.

Proof. To compute the above probability we will make use of Lemma 2.7 by conditioning on S_{n-1}. Then

$$\begin{aligned}
&\mathbb{P}(S_k \le tU_{n-1:k} \text{ for } k=1,...,n-1|S_{n-1}=s, S_n=t) \\
&\quad = \mathbb{P}(S_k \le tU_{n-1:k} \text{ for } k=1,...,n-1|S_{n-1}=s) \\
&\quad = \begin{cases} 1-\dfrac{s}{t}, & \text{if } 0<s\le t, \\ \quad 0, & \text{otherwise} \end{cases}
\end{aligned}$$

for $0 \le s \le t$. Hence, as $S_{n-1} \le S_n$, we have that

$$\begin{aligned}
&\mathbb{P}(S_k \le tU_{n-1:k} \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{E}\left[1-\frac{S_{n-1}}{t}\middle|S_n=t\right] \\
&\quad = 1-\frac{n-1}{n} \qquad \text{(by Lemma 2.6)}
\end{aligned}$$

which proves the result. $\diamondsuit$

Returning to the joint distribution of the length of a busy period and the number of customers served, we must, from (2.8), compute

$$\mathbb{P}(tU_{n-1:k} \le S_k \text{ for } k=1,...,n-1|S_n=t),$$

as $(T_1^*,...,T_{n-1}^*) = (tU_{n-1:1},...,tU_{n-1:n-1})$ for some i.i.d. uniform(0,1) variables $U_1,...,U_{n-1}$. Now, since $1-U \stackrel{d}{=} U$ for any uniform(0,1) variable U, it follows that

$$(1-U_{n-1:n-1},...,1-U_{n-1:1}) \;\stackrel{d}{=}\; (U_{n-1:1},...,U_{n-1:n-1}).$$

Hence, upon replacing $U_{n-1:k}$ by $1-U_{n-1:n-k}$ throughout, $1 \le k \le n-1$, we obtain

$$\begin{aligned}
&\mathbb{P}(tU_{n-1:k} \le S_k \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{P}(t-tU_{n-1:n-k} \le S_k \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{P}(t-tU_{n-1:n-k} \le t-(S_n-S_k) \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{P}(tU_{n-1:n-k} \ge S_n-S_k \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{P}(tU_{n-1:n-k} \ge S_{n-k} \text{ for } k=1,...,n-1|S_n=t) \\
&\quad = \mathbb{P}(tU_{n-1:k} \ge S_k \text{ for } k=1,...,n-1|S_n=t)
\end{aligned}$$

where the next-to-last equality follows because the conditional laws of $(X_1,...,X_n)$ and $(X_n,...,X_1)$ given $S_n=t$ coincide, and so any probability statement involving the X_i's and further random variables independent of $X_1,...,X_n$ remains valid if X_1 is replaced by X_n, X_2 by

X_{n-1},...,X_k by X_{n-k}, ...,X_n by X_1. Hence, we see that

$$\begin{aligned}\mathbb{P}(tU_{n-1:k} \leq S_k \text{ for } k = 1, ..., n-1 | S_n = t) \\ &= \mathbb{P}(tU_{n-1:k} \geq S_k \text{ for } k = 1, ..., n-1 | S_n = t) \\ &= \frac{1}{n} \qquad \text{(from Lemma 2.8)}.\end{aligned}$$

Now, from (2.6), if we let

$$B(t,n) \stackrel{\text{def}}{=} \mathbb{P}(\text{busy period is of length} \leq t, n \text{ customers served in a busy period}),$$

then

$$\frac{d}{dt}B(t,n) \;=\; e^{-\lambda t}\,\frac{(\lambda t)^{n-1}}{n!}\,G_n(dt)$$

or

$$B(t,n) \;=\; \int_0^t e^{-\lambda t}\,\frac{(\lambda t)^{n-1}}{n!}\;G_n(dt).$$

The distribution function of the length of a busy period, call it $B(t) \stackrel{\text{def}}{=} \sum_{n\geq 1} B(t,n)$, is then given by

$$B(t) \;=\; \sum_{n\geq 1}\int_0^t e^{-\lambda t}\,\frac{(\lambda t)^{n-1}}{n!}\;G_n(dt). \qquad \spadesuit$$

3. The Nonhomogeneous Poisson Process

We will now generalize the Poisson process by allowing the arrival rate to be time-dependent. Again we will provide a number of definitions that focus on different characterizing aspects of this type of process.

Definition 3.1. [The Axiomatic Way] A counting process $(N(t))_{t\geq 0}$ is said to be a *nonstationary* or *nonhomogeneous Poisson process with rate (or intensity) function* $\lambda(t)$, $t \geq 0$, if:

(NPP1) $N(0) = 0$.

(NPP2) The process has independent increments.

(NPP3) The number of events in any time interval $(s,t]$ is Poisson distributed with mean $\int_s^t \lambda(x)\,dx$, i.e. $N((s,t]) \stackrel{d}{=} Poi(m(t) - m(s))$, where $m(t) \stackrel{\text{def}}{=} \int_0^t \lambda(x)\,dx$ is the *cumulative rate function*.

Plainly, condition (NPP3) states that $(N(t))_{t\geq 0}$ does not have stationary increments unless $\lambda(t) \equiv \lambda$ for some $\lambda > 0$. It should further be understood from this condition that the rate function $\lambda(t)$ is supposed to be nonnegative and locally integrable, i.e. $\int_0^t \lambda(x)\,dx < \infty$ for all $t > 0$. Note that the function

$$m(t) \;=\; \int_0^t \lambda(x)\;dx, \quad t \geq 0$$

defines a locally finite measure ν on $[0, \infty)$ via

$$\nu((s,t]) \stackrel{\text{def}}{=} m(t) - m(s), \quad 0 \le s < t < \infty,$$

which is usually called the *intensity measure of* the process. In the homogeneous case $\lambda(t) \equiv \lambda > 0$ it obviously equals λ times Lebesgue measure on $[0, \infty)$.

Our second definition of a nonhomogeneous Poisson process provides an infinitesimal description and as such must impose the additional condition that the rate function be continuous. It is therefore more restrictive than the previous one.

Definition 3.2. [By Infinitesimal Description]. A counting process $(N(t))_{t \ge 0}$ is said to be a nonhomogeneous Poisson process with continuous rate function $\lambda(t)$, $t \ge 0$, if:

(NPP1) $N(0) = 0$.
(NPP2) The process has independent increments.
(NPP4) $\mathbb{P}(N(t+h) - N(t) = 1) = \lambda(t)h + o(h)$.
(NPP5) $\mathbb{P}(N(t+h) - N(t) \ge 2) = o(h)$.

As in the homogeneous case, it is straightforward to assess that processes defined by 3.1 (with continuous rate function) form a subclass of those defined by 3.2. For the converse, more has to be done but follows similar lines as in the homogeneous case.

Proof of "3.2 $\Rightarrow$ 3.1" when $\lambda(t)$ is continuous. Assuming (NPP1,2) and (NPP4,5), the task is to verify (NPP3), i.e. $N(s+t) - N(s) \stackrel{d}{=} Poi(m(s+t) - m(s))$ for each $s \ge 0$ and $t > 0$. Fix s and define

$$P_n(t) \stackrel{\text{def}}{=} \mathbb{P}(N(s+t) - N(s) = n), \quad n \in \mathbb{N}_0,$$

so that

$$P_n(t) \;=\; e^{-(m(s+t)-m(s))} \frac{[m(s+t) - m(s)]^n}{n!}, \quad n \in \mathbb{N}_0 \tag{3.1}$$

must be verified. Start by considering $P_0(t)$ for which a differential equation can be derived in the following manner: We leave it to the reader as an exercise to show that

$$\mathbb{P}(N(s+t) - N(s) = 0) > 0 \quad \text{for all } 0 \le s < t < \infty.$$

For $h > 0$, we infer with the help of (NPP2) and (NPP4,5) that

$$\begin{aligned} P_0(t+h) &= \mathbb{P}(N(s+t+h) - N(s) = 0) \\ &= \mathbb{P}(N(s+t) - N(s) = 0, N(s+t+h) - N(s+t) = 0) \\ &= \mathbb{P}(N(s+t) - N(s) = 0)\, \mathbb{P}(N(s+t+h) - N(s+t) = 0) \\ &= P_0(t)[1 - \lambda(s+t)h + o(h)] \end{aligned} \tag{3.2}$$

and thereupon

$$\lim_{h\downarrow 0} \frac{P_0(t+h)-P_0(t)}{h} \;=\; -\lambda(s+t)P_0(t).$$

Replacing s with $s-h$ in (3.2), we also have

$$P_0(t) \;=\; P_0(t-h)[1-\lambda(s+t-h)h+o(h)] \tag{3.3}$$

and thus see that $P_0(t-h)\to P_0(t)$ as $h\downarrow 0$. By combining this with the continuity of $\lambda(t)$, we further infer from (3.3) that

$$\lim_{h\downarrow 0} \frac{P_0(t-h)-P_0(t)}{-h} \;=\; \lim_{h\downarrow 0} -\lambda(s+t-h)P_0(t-h)+\frac{o(h)}{h} \;=\; -\lambda(s+t)P_0(t).$$

Consequently, $P_0(t)$ is differentiable and satisfies

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

or (recalling that $P_0(t)>0$ for all $t\geq 0$)

$$\log P_0(t) \;=\; -\int_0^t \lambda(s+u)\; du$$

or

$$P_0(t) \;=\; e^{-(m(s+t)-m(s))}.$$

The remainder of the verification of (3.1) follows similarly and is left as an exercise. $\diamondsuit$

A particularly quick way of introducing the nonhomogeneous Poisson process, which at the same time settles the question of its existence, is by changing the time scale of a standard Poisson process.

Definition 3.3. [The Constructive Way: Time Change]. A counting process $(N(t))_{t\geq 0}$ is said to be a nonhomogeneous Poisson process with rate function $\lambda(t)$, $t\geq 0$, if

$$N(t) \;=\; \widehat{N}(m(t)), \quad t\geq 0, \tag{3.4}$$

for a standard Poisson process $(\widehat{N}(t))_{t\geq 0}$.

The reader will readily check that Definition 3.3 implies Definition 3.1. Conversely, if $(N(t))_{t\geq 0}$ is a nonhomogeneous Poisson process with cumulative rate function $m(t)$, the standard Poisson process $(\widehat{N}(t))_{t\geq 0}$ in (3.4) can be obtained via a time change based on the pseudo-inverse $m^{-1}(t)$ of $m(t)$, defined as

$$m^{-1}(t) \;\stackrel{\text{def}}{=}\; \inf\{s\geq 0: m(s)\geq t\}, \quad t\geq 0. \tag{3.5}$$

The continuity of $m(t)$ implies $m(m^{-1}(t)) = t$, while $m^{-1}(m(t)) = t_{\min}$ with $t_{\min}$ being the minimal s such that $m(s) = m(t)$. Since $m^{-1}(t)$ is nondecreasing and

$$\widehat{N}((s,t]) \;=\; N((m^{-1}(s), m^{-1}(t)]) \;\stackrel{d}{=}\; Poi(m(m^{-1}(t)) - m(m^{-1}(s))) \;=\; Poi(t-s)$$

for all $0 \le s < t < \infty$, we infer that $\widehat{N}(t) \stackrel{\text{def}}{=} N(m^{-1}(t))$, $t \ge 0$, constitutes a standard Poisson process, and it obviously also satisfies (3.4).

Tagging revisited. Suppose we are given a homogeneous Poisson process $(N(t))_{t\ge 0}$ with rate $\lambda > 0$ and furthermore a rate function $\lambda(t)$, $t \ge 0$, which is bounded by λ. An arrival when occurring from $(N(t))_{t\ge 0}$ at time s is classified (tagged) as a type I event with probability $p(s) = \lambda(s)/\lambda$ and as a type II event with probability $1 - p(s)$. Let $(N_1(t))_{t\ge 0}$ and $(N_2(t))_{t\ge 0}$ denote the resulting counting processes of type I and type II events. By Proposition 2.3, $N_1(t)$ and $N_2(t)$ are independent Poisson variables with means $m(t)$ and $\lambda t - m(t)$, respectively, for each $t \ge 0$. The now natural question whether the whole processes $(N_1(t))_{t\ge 0}$ and $(N_2(t))_{t\ge 0}$ are independent Poisson processes is affirmatively answered by the following result.

Theorem 3.4. *In the situation just described the following assertions hold:*

(a) $(N_1(t))_{t\ge 0}$ *and* $(N_2(t))_{t\ge 0}$ *are nonhomogeneous Poisson processes with rate functions* $\lambda(t)$ *and* $\lambda - \lambda(t)$, *respectively.*

(b) $(N_1(t))_{t\ge 0}$ *and* $(N_2(t))_{t\ge 0}$ *are independent.*

PROOF. We restrict ourselves to some intuitive explanations. Given $n \ge 2$ pairwise disjoint time intervals, arrivals within these intervals from the proposal process $(N(t))_{t\ge 0}$ occur independently. Moreover, since each classification result depends on the proposal process only through the particular pertinent arrival that is to be classified, we see that the number of type I as well as type II events within these time intervals are also independent. But this in combination with what has been inferred from Proposition 2.1 easily shows the theorem. $\diamondsuit$

Remarks. (a) If $\lambda(t)$ is merely locally bounded the previous result is still valid in the following local sense: For any fixed $T > 0$, let $\lambda = \lambda_T > 0$ be such that $\sup_{0\le t\le T} \lambda(t) \le \lambda$. Then Theorem 3.4 remains true when sampling from a homogeneous Poisson process with rate λ and using the same classification procedure but restricted to the time interval $[0, T]$.

(b) One may also generalize Theorem 3.4 — and Proposition 2.1 as well — by allowing more than two types of events. More precisely, let $m \ge 2$ and $\lambda_1(t), ..., \lambda_m(t)$ be rate functions such that $\sum_{j=1}^m \lambda_j(t) \equiv \lambda > 0$. Any event of a given rate λ Poisson process is classified as a type k event with probability $\lambda_k(t)/\lambda$ when occurring at time t. Let $(N_k(t))_{t\ge 0}$ be the resulting counting process of type k events for $k = 1, ..., m$. Then the conclusion is that $(N_1(t))_{t\ge 0}, ..., (N_m(t))_{t\ge 0}$ are independent nonhomogeneous Poisson processes with rate functions $\lambda_1(t), ..., \lambda_m(t)$, respectively.

Example 3.5. ***The Output Process of an Infinite Server Poisson Queue.*** It turns out that the output process of the $M/G/\infty$ queue — that is of the infinite server queue having Poisson arrivals and general service distribution G (☞ Example 2.4) — is a nonhomogeneous Poisson process having intensity function $\lambda(t) = \lambda G(t)$. To prove this we shall argue that

(1) the number of departures in $(s, s+t]$ is Poisson distributed with mean $\lambda \int_s^{s+t} G(x)\, dx$, and

(2) the numbers of departures in disjoint time intervals are independent.

To prove statement (1), call an arrival type I if it departs in the interval $(s, s+t]$. Then an arrival at x will be type I with probability

$$p(y) \stackrel{\text{def}}{=} \begin{cases} G(s+t-y) - G(s-y), & \text{if } y < s, \\ G(s+t-y), & \text{if } s < y \le s+t, \\ 0, & \text{if } y > s+t. \end{cases}$$

Hence, from Proposition 2.1 the number of such departures will be Poisson distributed with mean

$$\begin{aligned} \lambda \int_0^{s+t} p(y)\ dy \ &= \ \lambda \int_0^s \big(G(s+t-y) - G(s-y)\big)\ dy \ + \ \lambda \int_s^{s+t} G(s+t-y)\ dy \\ &= \ \lambda \int_s^{s+t} G(y)\ dy. \end{aligned}$$

To prove statement (2), let $I_1, ..., I_n$, $n \ge 2$, be pairwise disjoint time intervals of the form $(a, b]$ and call an arrival type k if it departs in I_k for $k = 1, ..., n$, and call it type $n+1$ otherwise. Again, from Proposition 2.1 (and Remark (b) above), it follows that the number of departures in $I_1, ..., I_n$ are mutually independent Poisson variables.

Using statements (1) and (2) it is clear that the output (departure) process $(N(t))_{t\ge 0}$, say, satisfies (NPP1-3) with $\lambda(t) = \lambda G(t)$, $t \ge 0$. ♠

Example 3.6. ***Record Values.*** Let $X_1, X_2, ...$ denote a sequence of i.i.d. nonnegative random variables with distribution function F, density function f and hazard rate function $\lambda(t) = f(t)/\overline{F}(t)$, where $\overline{F} \stackrel{\text{def}}{=} 1 - F$. Recall that $\lambda(t) = -(\log \overline{F})'(t) = -\overline{F}'(t)/\overline{F}(t)$ and therefore

$$\overline{F}(t) \ = \ \exp\left(- \int_0^t \lambda(s)\ ds \right), \quad t > 0.$$

We say that a *record* occurs at time n if $X_n > \max(X_1, ..., X_{n-1})$, where $X_0 \stackrel{\text{def}}{=} 0$. In this case X_n is called a *record value*. We claim that, if $N(t)$ denotes the number of record values less than or equal to t, then $(N(t))_{t\ge 0}$ is a nonhomogeneous Poisson process with rate function $\lambda(t)$, $t \ge 0$.

In order to prove this claim we first give a formal definition of $N(t)$. Define $T_0 \stackrel{\text{def}}{=} 0$ and then recursively

$$T_n \stackrel{\text{def}}{=} \inf\{k > T_{n-1} : X_k > X_{T_{n-1}}\}, \quad n \geq 1.$$

Plainly, $T_1, T_2, ...$ are the times at which records occur, called *record epochs*, and $X_{T_1}, X_{T_2}, ...$ are the record values. Now

$$N(t) \stackrel{\text{def}}{=} \sum_{n\geq 1} \mathbf{1}_{(0,t]}(X_{T_n}), \quad t \geq 0. \tag{3.6}$$

In the following, we first consider the case where the X_n have a standard exponential distribution, i.e. $\overline{F}(t) = e^{-t}$ or $\lambda(t) = 1$ for all $t \geq 0$. We claim that the $D_n \stackrel{\text{def}}{=} X_{T_n} - X_{T_{n-1}}$, $n \geq 1$, are also i.i.d. standard exponentials and verify this by an induction on n:

For $n = 0$ it suffices to note that $D_1 = X_{T_1} - X_{T_0} = X_1$. Assume now that $D_1, ..., D_n$ are i.i.d. standard exponentials (inductive hypothesis). We must verify that D_{n+1} is independent of $D_1, ..., D_n$ and also a standard exponential. Put $\tau \stackrel{\text{def}}{=} T_{n+1} - T_n = \inf\{k \geq 1 : X_{T_n+k} > X_{T_n}\}$. Lemma 3.6 below shows that the sequence $X_{T_n+1}, X_{T_n+2}, ...$ forms a copy of $X_1, X_2, ...$ and is further independent of T_n and $D_1, ..., D_n$. By using this fact, we infer for any $k \geq 1$ and $x_1, ..., x_n, t > 0$ (treating densities like probabilities)

$$\begin{aligned}
&\mathbb{P}(D_{n+1} = t, \tau = k, D_1 = x_1, ..., D_n = x_n) \\
&\quad = \mathbb{P}(X_{T_n+k} - X_{T_n} = t, \tau = k, D_1 = x_1, ..., D_n = x_n) \\
&\quad = \mathbb{P}(X_{T_n+k} = t + s_n, \tau = k, D_1 = x_1, ..., D_n = x_n) \qquad [s_n \stackrel{\text{def}}{=} x_1 + ... + x_n] \\
&\quad = \mathbb{P}(X_{T_n+k} = t + s_n, X_{T_n+j} \leq s_n, 1 \leq j < k, D_1 = x_1, ..., D_n = x_n) \\
&\quad = e^{-t-s_n}(1 - e^{-s_n})^{k-1}\mathbb{P}(D_1 = x_1, ..., D_n = x_n) \\
&\quad = e^{-t-s_n}(1 - e^{-s_n})^{k-1}\mathbb{P}(D_1 = x_1) \cdot ... \cdot \mathbb{P}(D_n = x_n) \qquad [\text{inductive hypothesis}]
\end{aligned}$$

and thereby upon summation over k

$$\begin{aligned}
&\mathbb{P}(D_{n+1} = t, D_1 = x_1, ..., D_n = x_n) \\
&\quad = \sum_{k\geq 1} \mathbb{P}(D_{n+1} = t, \tau = k, D_1 = x_1, ..., D_n = x_n) \\
&\quad = e^{-t-s_n}\, \mathbb{P}(D_1 = x_1) \cdot ... \cdot \mathbb{P}(D_n = x_n) \sum_{k\geq 1}(1 - e^{-s_n})^{k-1} \\
&\quad = e^{-t}\, \mathbb{P}(D_1 = x_1) \cdot ... \cdot \mathbb{P}(D_n = x_n).
\end{aligned}$$

This clearly proves the assertion.

Having shown that the X_{T_n} are sums of i.i.d. standard exponentials, we directly conclude from Definition 1.3 that $(N(t))_{t\geq 0}$ forms a standard Poisson process because, by (3.6), the X_{T_n} are the jump epochs of this counting process.

Turning to the general situation, we first note that, since F is continuous, the sequence $F(X_1), F(X_2), ...$ consists of i.i.d. uniform(0,1) random variables. Indeed, for $t \in (0, 1)$ put

$u \stackrel{\text{def}}{=} \max\{u : F(u) = t\}$ which exists by the continuity of F. Then

$$\mathbb{P}(F(X_n) \le t) \;=\; \mathbb{P}(X_n \le u) \;=\; F(u) \;=\; t$$

shows that X_n has the asserted distribution. Next put

$$Y_n \;\stackrel{\text{def}}{=}\; -\log \overline{F}(X_n) \;=\; \int_0^{X_n} \lambda(s)\, ds \;=\; m(X_n)$$

for $n \ge 1$ and check that these variables are i.i.d. standard exponentials. By the strict monotonicity of the transformation $t \mapsto -\log(1-t)$, the T_n are also the record epochs for the sequence $(Y_n)_{n\ge1}$. Consequently, the counting process

$$\widehat{N}(t) \;\stackrel{\text{def}}{=}\; \sum_{n\ge1} \mathbf{1}_{(0,t]}(Y_{T_n}), \quad t \ge 0$$

constitutes a standard Poisson process by what has been proved in the first step. Finally, as

$$N(t) \;=\; \sum_{n\ge1} \mathbf{1}_{(0,-\log \overline{F}(t)]}(Y_{T_n}) \;=\; \sum_{n\ge1} \mathbf{1}_{(0,m(t)]}(Y_{T_n}) \;=\; \widehat{N}(m(t))$$

for all $t \ge 0$, we arrive at the asserted result by an appeal to Definition 3.3. ♠