Notes on Exponential Distribution and Poisson Process

Yangxinyu Xie

Summer 2019

Last updated January 11, 2020

1. Exponential Distribution

❖

1.1 Definition ❖

Definition 1.1. We say that a continuous random variables x has an **exponential distribution** if its probability density function is given by

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1.1.1}$$

We write $x \sim exp(\lambda)$.

The cumulative distribution function is thus given by

$$F(x) = \int_{-\infty}^{x} f(y)dy = \begin{cases} 1 - e^{-\lambda x} & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{1.1.2}$$

The moment generating function is then

$$\phi(t) = \mathbb{E}(e^{tx}) = \int_0^\infty e^{tx}\lambda e^{-\lambda x}dx = \frac{\lambda}{\lambda - t} \tag{1.1.3}$$

The expectation can be obtained by

$$\mathbb{E}(x) = \left.\frac{d\phi(t)}{dt}\right|_{t=0} = \left.\frac{\lambda}{(\lambda - t)^2}\right|_{t=0} = \frac{1}{\lambda} \tag{1.1.4}$$

and the variance

$$V(x) = \mathbb{E}(x^2) - \mathbb{E}(x)^2 = \left.\frac{d^2\phi(t)}{d^2t}\right|_{t=0} - \frac{1}{\lambda^2} = \left.\frac{2\lambda}{(\lambda - t)^3}\right|_{t=0} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \tag{1.1.5}$$

1.2 Memorylessness ❖

Definition 1.2. We say that a random variable X is **memoryless** if

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s) \text{ for all } s, t \geq 0 \tag{1.2.1}$$

or equivalently by Bayes' theorem,

$$\mathbb{P}(X > s + t) = \mathbb{P}(X > s)\mathbb{P}(X > t) \text{ for all } s, t \geq 0 \tag{1.2.2}$$

Theorem 1.3. *An exponential random variable X is memoryless.*

Proof. We can verify the memoryless property directly from the definition.

$$\mathbb{P}(X > s + t) = 1 - F(s + t) = e^{-\lambda(s+t)} = e^{-\lambda s}e^{-\lambda t} = (1 - F(s))(1 - F(t)) = \mathbb{P}(X > s)\mathbb{P}(X > t) \tag{1.2.3}$$

□

Example 1.4. Suppose the demand for a certain commodity is exponentially distributed with rate λ. If the commodity costs the store c per unit, while it can be sold for $s > c$ per unit, and the leftover units will be worthless at the end of the month, we are interested how many units the store should purchase to maximise its expected profit $\mathbb{E}(P)$.

1. Let X be the demand for this commodity and t be the number of units the store purchased. Hence, we have
$$P = s\min(X,t) - ct = s[X - (X-t)^+] - ct \tag{1.2.4}$$
where $(X-t)^+ = X - t$ if $X \geq t$, and $(X-t)^+ = 0$ otherwise. Hence,
$$\begin{aligned}\mathbb{E}[(X-t)^+] &= \mathbb{E}[X-t|X\geq t]\mathbb{P}(X\geq t) + \mathbb{E}[0|X<t]\times\mathbb{P}(X<t)\\ &= \mathbb{E}[X-t|X\geq t]\mathbb{P}(X\geq t)\\ &= \frac{e^{-\lambda t}}{\lambda} \text{ memoryless property}\end{aligned} \tag{1.2.5}$$
Therefore,
$$\mathbb{E}[\min(X,t)] = \frac{1}{\lambda} - \frac{e^{-\lambda t}}{\lambda} \tag{1.2.6}$$
and
$$\mathbb{E}(P) = \frac{s}{\lambda} - \frac{se^{-\lambda t}}{\lambda} - ct \tag{1.2.7}$$
Setting the derivative to 0,
$$\frac{d\mathbb{E}(P)}{dt} = se^{-\lambda t} - c = 0 \tag{1.2.8}$$
We obtain
$$t = \frac{\log(s) - \log(c)}{\lambda} \tag{1.2.9}$$

2. Now let's suppose that each unit left at the end of the month can be returned for a value $r < c$ per unit. Then,
$$\mathbb{E}(P) = \frac{s}{\lambda} - \frac{se^{-\lambda t}}{\lambda} - ct + r\mathbb{E}[(t-X)^+] \tag{1.2.10}$$
where
$$\mathbb{E}[(t-X)^+] = \mathbb{E}[t - \min(X,t)] = t - \mathbb{E}[\min(X,t)] \tag{1.2.11}$$
Therefore,
$$\mathbb{E}(P) = (s-r)(\frac{1}{\lambda} - \frac{e^{-\lambda t}}{\lambda}) - (c-r)t \tag{1.2.12}$$
Setting the derivative to 0,
$$\frac{d\mathbb{E}(P)}{dt} = (s-r)e^{-\lambda t} - (c-r) = 0 \tag{1.2.13}$$
Therefore,
$$t = \frac{\log(s-r) - \log(c-r)}{\lambda} \tag{1.2.14}$$

Definition 1.5. Suppose X is a continuous positive random variable with cumulative distribution function F and probability density function f. The **failure(or hazard) rate function** $r(t)$ is defined by
$$r(t) = \frac{f(t)}{1-F(t)} \tag{1.2.15}$$

To interpret failure rate function, consider a machine that has been used for a time period t, the probability that the machine will fail within the additional time dt is then
$$\mathbb{P}(X < t+dt|X\geq t) = \frac{\mathbb{P}(t\leq X<t+dt)}{\mathbb{P}(X\geq t)} = \frac{f(t)dt}{1-F(t)} \approx r(t)dt \tag{1.2.16}$$
From this view, we can interpret $r(t)$ as the conditional probability density function that a machine will fail given that it has been used for some time t. Note that the failure rate function uniquely defines the distribution. Too see this, we have
$$r(t) = \frac{\frac{d}{dt}F(t)}{1-F(t)} \tag{1.2.17}$$

Integrating on both sides, we have

$$\int_0^t r(t)dt + C = -\log(1-F(t)) \tag{1.2.18}$$

Letting $t = 0$, we have $0 + C = 0$. Therefore, we can write the cumulative distribution function as

$$F(t) = 1 - e^{-\int_0^t r(t)dt} \tag{1.2.19}$$

If the distribution is memoryless, that is

$$(1-F(a))(1-F(b)) = e^{-\int_0^a r(t)dt}e^{-\int_0^b r(t)dt} = e^{-\int_0^{a+b} r(t)dt} = (1-F(a+b)) \tag{1.2.20}$$

then $r(t)$ must be some constant. Hence, $1 - F(t) = e^{-ct}$ for some constant c. In fact, in the case of exponential random variable, we have

$$r(t) = \frac{f(t)}{1-F(t)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda \tag{1.2.21}$$

This implies that exponential distribution is the only distribution that is memoryless.

Definition 1.6. Suppose we have n independent exponential distributions with respective rates $\lambda_1, \lambda_2, ...\lambda_n$. Suppose we also have a discrete random variable T such that $\mathbb{P}(T = i) = P_i$ and $\sum_{i=1}^n P_i = 1$. We say that the random variable X_T is a **hyperexponential** random variable.

Specifically, the cumulative distribution function of a hyperexponential random variable is

$$F(t) = 1 - \mathbb{P}(X_T > t) = 1 - \sum_{i=1}^n \mathbb{P}(X_T > t|T = i)\mathbb{P}(T = i) = 1 - \sum_{i=1}^n \exp(-\lambda_i t)\mathbb{P}(T = i) \tag{1.2.22}$$

and the probability density function is

$$f(t) = \sum_{i=1}^n \lambda_i \exp(-\lambda_i t)\mathbb{P}(T = i) \tag{1.2.23}$$

The failure rate function is then

$$r(t) = \frac{f(t)}{1-F(t)} = \frac{\sum_{j=1}^n \lambda_j \exp(-\lambda_j t)\mathbb{P}(T = j)}{\sum_{i=1}^n \exp(-\lambda_i t)\mathbb{P}(T = i)} = \sum_{j=1}^n \lambda_j \mathbb{P}(T = j|X_T > t) \tag{1.2.24}$$

Especially, if $\lambda_1 < \lambda_i$ for all i,

$$\lim_{t\to\infty} \mathbb{P}(T = 1|X_T > t) = \lim_{t\to\infty} \frac{\exp(-\lambda_1 t)\mathbb{P}(T = 1)}{\exp(-\lambda_1 t)\mathbb{P}(T = 1) + \sum_{i=2}^n \exp(-\lambda_i t)\mathbb{P}(T = i)} = 1 \tag{1.2.25}$$

The intuition is that the longer the machine runs, the higher the probability that this machines has a lower failure rate is. Therefore, we have $\lim_{t\to\infty} r(t) = \min_i \lambda_i$.

1.3 Gamma Distribution ❖

Definition 1.7. We say that a continuous random variable x with density function given by

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x}(\lambda x)^{\alpha-1}}{\Gamma(\alpha)}, & \text{if } x \geq 0 \\ 0, & \text{if } x < 0 \end{cases} \tag{1.3.1}$$

for some $\lambda > 0, \alpha > 0$ is a **gamma random variable** with parameters α, λ where

$$\Gamma(\alpha) = \int_0^\infty e^{-x}x^{\alpha-1}dx \tag{1.3.2}$$

We write $x \sim gamma(\alpha, \lambda)$.

$\Gamma(\alpha)$ is often referred to as the **gamma function**. We can see that

$$\begin{aligned}\Gamma(n+1) &= \int_0^\infty e^{-x}x^n dx \\ &= -e^{-x}x^n\Big|_0^\infty + n\int_0^\infty e^{-x}x^{n-1}dx \\ &= n\int_0^\infty e^{-x}x^{n-1}dx\end{aligned} \tag{1.3.3}$$

Thus, we can prove by induction that $\Gamma(n) = (n-1)!$.

Proposition 1.8. *Suppose $X_1, ..., X_n$ are independent exponential random variables with the same rate λ, then $X_1 + X_2 + \ldots + X_n \sim gamma(n, \lambda)$.*

Proof. We can easily see that if $X \sim exp(\lambda)$, then $X \sim gamma(1, \lambda)$. In fact, we can further suppose that $X_1, X_2, \ldots$ are independently exponentially distributed with the same rate λ. Assuming that $X_1 + X_2 + \ldots + X_{n-1} \sim gamma(n-1, \lambda)$, then

$$\begin{aligned}f_{X_1+X_2+\ldots+X_n}(t) &= \int_0^t f_{X_n}(t-s)f_{X_1+X_2+\ldots+X_{n-1}}(s)ds \\ &= \int_0^t \lambda e^{-\lambda(t-s)}\frac{\lambda e^{-\lambda s}(\lambda s)^{n-1-1}}{\Gamma(n-1)}ds \\ &= \lambda e^{-\lambda t}\int_0^t \frac{\lambda^{n-1}s^{n-2}}{\Gamma(n-1)}ds \\ &= \frac{\lambda e^{-\lambda t}(\lambda s)^{n-1}}{\Gamma(n)}\end{aligned} \tag{1.3.4}$$

Hence, we prove by induction that $X_1 + X_2 + \ldots + X_n \sim gamma(n, \lambda)$. □

1.4 Minimum Random Variable ❖

Proposition 1.9. *Suppose $X_1, ..., X_n$ are independent exponential random variables with respective rates $\lambda_1, ..., \lambda_n$, then $\min_i X_i \sim exp(\sum_{i=1}^n \lambda_i)$.*

Proof. By independence, we obtain

$$\mathbb{P}(\min_i X_i > x) = \prod_{i=1}^n \mathbb{P}(X_i > x) = \prod_{i=1}^n e^{-\lambda_i x} = \exp(-x \cdot \sum_{i=1}^n \lambda_i) \tag{1.4.1}$$

□

We now consider two independent exponential random variables $X_1 \sim exp(\lambda_1)$ and $X_2 \sim exp(\lambda_2)$ and are interested in the probability that $X_1 < X_2$, then

$$\begin{aligned}\mathbb{P}(X_1 < X_2) &= \int_0^\infty \mathbb{P}(X_1 < X_2|X_1 = x)\lambda_1 e^{-\lambda_1 x}dx \\ &= \int_0^\infty \mathbb{P}(x < X_2)\lambda_1 e^{-\lambda_1 x}dx \\ &= \int_0^\infty e^{-\lambda_2 x}\lambda_1 e^{-\lambda_1 x}dx \\ &= \frac{\lambda_1}{\lambda_1 + \lambda_2}\end{aligned} \tag{1.4.2}$$

Suppose $X_1, ..., X_n$ are independent exponential random variables with respective rates $\lambda_1, ..., \lambda_n$, $\lambda_i \neq \lambda_j$ if $i \neq j$, then we can generalise the preceding result 1.4.2 to the following

$$\mathbb{P}(X_i = \min_j X_j) = \frac{\lambda_i}{\sum_{j=1}^n \lambda_j} \tag{1.4.3}$$

Example 1.10. Suppose there are n cells, each of whom has a weight w_i. These cells will die in a random order. Specifically, suppose S is the set of remaining cells. The probability that cell i will die next is given by $w_i/(\sum_{j\in S} w_j)$. Let A be the number of cells alive at the moment when cell $1, \dots, l$ all died and we are interested in $\mathbb{E}(A)$. Let X_i be a lifetime of the cell i, then

$$\mathbb{P}(X_i \leq X_j \text{ for all } j \in S, j \neq i) = \frac{w_i}{\sum_{j\in S} w_j} \tag{1.4.4}$$

From equation 1.4.3 we can assume that each X_i is exponentially distributed with rate w_i. Let A_j denote the event that the jth cell is alive. Hence $A = \sum_{j=k+1}^n A_j$ and for each A_j

$$\begin{aligned}
\mathbb{E}(A_j) &= \int_0^\infty \mathbb{P}(X_j > \max_{i\in\{1,\dots,k\}} X_i | X_j = x) w_j e^{-w_j x} dx \\
&= \int_0^\infty \mathbb{P}(X_i < x, i \in \{1,\dots,k\}) w_j e^{-w_j x} dx \\
&= \int_0^\infty \prod_{i=1}^k (1 - e^{-w_i x}) w_j e^{-w_j x} dx \\
&= \int_0^1 \prod_{i=1}^k (1 - y^{w_i/w_j}) dy \text{ where } y = e^{-w_j x}
\end{aligned} \tag{1.4.5}$$

Proposition 1.11. *Suppose $X_1, ..., X_n$ are independent exponential random variables with respective rates $\lambda_1, ..., \lambda_n$, $\lambda_i \neq \lambda_j$ if $i \neq j$. $\min_i X_i$ and the rank order of the variables $X_1, ..., X_n$ are independent.*

Proof. Suppose there is a sequence of the indices $i_1, ..., i_n$ such that $X_{i_1} < \ldots < X_{i_n}$. Using the memoryless property, we have

$$\mathbb{P}(X_{i_1} < \ldots < X_{i_n} | \min_i X_i > t) = \mathbb{P}(X_{i_1} - t < \ldots < X_{i_n} - t | \min_i X_i > t) = \mathbb{P}(X_{i_1} < \ldots < X_{i_n}) \tag{1.4.6}$$

□

1.5 Convolutions ❖

Definition 1.12. Suppose $X_1, ..., X_n$ are independent exponential random variables with respective rates $\lambda_1, ..., \lambda_n$, $\lambda_i \neq \lambda_j$ if $i \neq j$.The random variable $\sum_{i=1}^n X_i$ is said to be a **hypoexponential** random variable.

To obtain the probability density function of a hypoexponential random variable, we first consider $X_1 + X_2$.

$$\begin{aligned}
f_{X_1+X_2}(t) &= \int_0^t f_{X_1}(s) f_{X_2}(t-s) ds \\
&= \int_0^t \lambda_1 e^{-\lambda_1 s} \lambda_2 e^{-\lambda_2 (t-s)} ds \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 t} \int_0^t e^{(\lambda_2 - \lambda_1)s} ds \\
&= \lambda_1 \lambda_2 e^{-\lambda_2 t} \left[\frac{e^{(\lambda_2-\lambda_1)s}}{\lambda_2 - \lambda_1} \right]_0^t ds \\
&= \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_1 t} - \frac{\lambda_1\lambda_2}{\lambda_2 - \lambda_1} e^{-\lambda_2 t} \\
&= \frac{\lambda_2}{\lambda_2 - \lambda_1} \cdot \lambda_1 e^{-\lambda_1 t} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 t}
\end{aligned} \tag{1.5.1}$$

Following the preceding, we have for $X_1 + X_2 + X_3$,

$$\begin{aligned} f_{X_1+X_2+X_3}(t) &= \int_0^t f_{X_1+X_2}(s) f_{X_3}(t-s) ds \\ &= \int_0^t \Big(\frac{\lambda_2}{\lambda_2 - \lambda_1} \cdot \lambda_1 e^{-\lambda_1 s} + \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_2 s} \Big) \lambda_3 e^{-\lambda_3 (t-s)} ds \\ &= \lambda_3 e^{-\lambda_3 t} \int_0^t \frac{\lambda_2}{\lambda_2 - \lambda_1} \cdot \lambda_1 e^{-\lambda_3 - \lambda_1 s} ds + \lambda_3 e^{-\lambda_3 t} \int_0^t \frac{\lambda_1}{\lambda_1 - \lambda_2} \lambda_2 e^{-\lambda_3 - \lambda_2 s} \lambda_3 e^{-\lambda_3 (t-s)} ds \\ &= \sum_{i=1}^{3} \left[\lambda_i e^{-\lambda_i t} \cdot \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right] \end{aligned} \tag{1.5.2}$$

In fact, it can be shown by induction that

$$f_S(t) = \sum_{i=1}^{n} \left[\lambda_i e^{-\lambda_i t} \cdot \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right] \tag{1.5.3}$$

where $S = \sum_{i=1}^{n} X_i$. The cumulative probability function is thus

$$F_S(t) = 1 - \sum_{i=1}^{n} \left[e^{-\lambda_i t} \cdot \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right] \tag{1.5.4}$$

The failure rate function is then given by

$$r_S(t) = \frac{f_S(t)}{1 - F_S(t)} = \frac{\sum_{i=1}^{n} \left[\lambda_i e^{-\lambda_i t} \cdot \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right]}{\sum_{i=1}^{n} \left[e^{-\lambda_i t} \cdot \prod_{j \neq i} \frac{\lambda_j}{\lambda_j - \lambda_i} \right]} \tag{1.5.5}$$

Similarly, we have $\lim_{t \to \infty} r_S(t) = \min_i \lambda_i$.

2. Stationary Poisson Process

2.1 Counting Process

Definition 2.1. A stochastic process $\{N(t), t \geq 0\}$ is said to be a **counting process** if $N(t)$ represents the total number of events that occur by time t. Specifically, the counting process must satisfy

1. $N(t) \geq 0$,
2. $N(t) \in \mathbb{N}$,
3. If $s < t$, then $N(s) \leq N(t)$, and
4. For $s < t, N(t) - N(s)$ denotes the number of events that occur within the time interval $(s, t]$.

Definition 2.2. A counting process is said to possess **independent increments** if the number of events that occur in disjoint time intervals are independent.

Definition 2.3. A counting process is said to possess **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.

2.2 Poisson Process ❖

Definition 2.4. The function f is said to be $o(h)$ if

$$\lim_{h \to 0} \frac{f(h)}{h} = 0 \tag{2.2.1}$$

Definition 2.5. A **Stationary Poisson Process** with intensity λ is a counting process $N = \{N(t) : t \geq 0\}$ taking values in $S = \{0, 1, 2, \ldots\}$ such that

1. $N(0) = 0$.
2. $$\mathbb{P}(N(t+h) = m + n | N(t) = n) = \begin{cases} \lambda h + o(h), & \text{if } m = 1; \\ o(h), & \text{if } m > 1; \\ 1 - \lambda h + o(h), & \text{if } m = 0. \end{cases} \tag{2.2.2}$$
3. It possesses independent and stationary increments.

Theorem 2.6. *$N(t)$ has the Poisson distribution with mean λt; that is,*

$$\mathbb{P}(N(t) = j) = \frac{(\lambda t)^j}{j!} e^{-\lambda t} \tag{2.2.3}$$

Proof. By definition, we have if $j \neq 0$

$$\begin{aligned} \mathbb{P}(N(t+h) = j) &= \sum_{i=0}^{j} \mathbb{P}(N(t+h) = j | N(t) = i) \mathbb{P}(N(t) = i) \\ &= \lambda h \mathbb{P}(N(t) = j-1) + (1 - \lambda h) \mathbb{P}(N(t) = j) + o(h) \end{aligned} \tag{2.2.4}$$

Let $p_j(t)$ denote $\mathbb{P}(N(t) = j)$, we have

$$\begin{cases} p_j(t+h) = \lambda h p_{j-1}(t) + (1 - \lambda h) p_j(t) + o(h) & \text{if } j \neq 0, \\ p_0(t+h) = (1 - \lambda h) p_0(t) + o(h) & \text{if } j = 0. \end{cases} \tag{2.2.5}$$

Thus,

$$\begin{cases} p_j'(t+h) = \lim_{h \downarrow 0} \frac{p_j(t+h) - p_j(t)}{h} = \lambda p_{j-1}(t) - \lambda p_j(t) & \text{if } j \neq 0, \\ p_0'(t+h) = \lim_{h \downarrow 0} \frac{p_0(t+h) - p_0(t)}{h} = -\lambda p_0(t) & \text{if } j = 0. \end{cases} \tag{2.2.6}$$

with the boundary conditions

$$p_j(0) = \begin{cases} 0 & \text{if } j \neq 0, \\ 1 & \text{if } j = 0. \end{cases} \tag{2.2.7}$$

We define the generating function

$$G(s,t) = \mathbb{E}[s^{N(t)}] = \sum_{j=0}^{\infty} p_j(t) s^j \tag{2.2.8}$$

Therefore,

$$\begin{aligned} \frac{\partial G}{\partial t} &= \sum_{j=0}^{\infty} p_j'(t) s^j \\ &= \lambda + \lambda \sum_{j=1}^{\infty} \left[p_{j-1}(t) s^j - p_j(t) s^j \right] \\ &= \lambda(s-1)(1 + p_1(t)s + p_2(t)s^2 + \ldots) \\ &= \lambda(s-1)G \end{aligned} \tag{2.2.9}$$

Hence, we must have

$$G(s,t) = e^{\lambda(s-1)t} = e^{-\lambda t}\sum_{j=0}^{\infty}\frac{(\lambda t)^j}{j!}s^j \tag{2.2.10}$$

which concludes the proof. □

2.3 Interarrival and Waiting Time Distributions ❖

Definition 2.7. For a given Poisson process, we let T_n denote the elapsed time between occurrence of the $(n-1)$th and the nth event. The sequence $\{T_n, n = 1, 2, \ldots\}$ is called the **sequence of interarrival times**.

Proposition 2.8. *Each $T_n, n = 1, 2, \ldots$ are independent identically distributed exponential random variable with mean $1/\lambda$.*

Proof. Out of independent and stationary increments, we have for each n,

$$\mathbb{P}(T_n > t) = \mathbb{P}(N(t) = 0) = e^{-\lambda t} \tag{2.3.1}$$

□

Definition 2.9. For a given Poisson process, we let T_n denote the elapsed time between occurrence of the $(n-1)$th and the nth event. The random variable $\{S_n = \sum_{i=1}^n T_n, n = 1, 2, \ldots\}$ is called the **waiting time** until the nth event.

By Proposition 1.8 and 2.8, we have that $S_n \sim \mathit{gamma}(n, \lambda)$. That is

$$f_{S_n}(t) = \lambda e^{-\lambda t}\frac{(\lambda t)^{n-1}}{(n-1)!} \tag{2.3.2}$$

2.4 Event Types ❖

Given a Poisson process $N = \{N(t), t \geq 0\}$ that each time produces a type I event with probability p and a type II event with probability $(1-p)$. Let $N_1(t)$ and $N_2(t)$ denote respectively the number of type I and type II events occurring in $[0, t]$.

Proposition 2.10. *$N_1 = \{N_1(t), t \geq 0\}$ and $N_2 = \{N_2(t), t \geq 0\}$ are both Poisson processes with respective rates λp and $\lambda(1-p)$. Furthermore, the two poisson processes are independent.*

Proof. We first focus on N_1. It is easy to see the following from Definition 2.5

1. $N_1(0) = N(0) = 0$
2. $$\begin{aligned}\mathbb{P}(N_1(h) = 1) &= \mathbb{P}(N_1(h) = 1|N(h) = 1)\mathbb{P}(N(h) = 1) + \mathbb{P}(N_1(h) = 1|N(h) > 1)\mathbb{P}(N(h) > 1)\\ &= p(\lambda h + o(h)) + o(h) = \lambda p h + o(h)\end{aligned} \tag{2.4.1}$$

   and

   $$\mathbb{P}(N_1(h) > 1) \leq \mathbb{P}(N_1(h) = 1|N(h) > 1)\mathbb{P}(N(h) > 1) = o(h) \tag{2.4.2}$$
3. N_1 inherits independent and stationary increments from N.

The same can be obtained for N_2. To see the independence of these two random variables, we use Theorem 2.6

$$\begin{aligned}\mathbb{P}(N_1(t) = m, N_2(t) = n) &= \mathbb{P}(N_1(t) = m, N_2(t) = n|N(t) = m+n)\mathbb{P}(N(t) = m+n)\\ &= \binom{m+n}{m}p^m(1-p)^n\frac{(\lambda t)^{m+n}}{(m+n)!}e^{-\lambda t}\\ &= \frac{(\lambda p t)^m}{m!}e^{-\lambda p t}\cdot\frac{[\lambda(1-p)t]^n}{n!}e^{-\lambda(1-p)t}\\ &= \mathbb{P}(N_1(t) = m)\mathbb{P}(N_2(t) = n)\end{aligned} \tag{2.4.3}$$

□

It is easy to generalise Proposition 2.10 into $m, m \geq 2$ types of events with respective rate $\lambda p_i, 1 \leq i \leq m$. We use the following example to show an application of the preceding result.

Example 2.11 (**The Coupon Collecting Problem).** The classical coupon collection problem states that we aim to collect m types of coupon by each time drawing a coupon independently, which will be type $i, 1 \leq i \leq m$ with probability $p_i, \sum_{i=1}^m p_i = 1$. Let N denote the number of coupons we need to collect in order to complete a collection and we look for the expectation $\mathbb{E}[N]$.
In fact, we can view this problem as a Poisson process $N = \{N(t) : t \geq 0\}$ with rate $\lambda = 1$. Let $N_j(t)$ be the number of type j coupons collected by time t. Then $N_j = \{N_j(t) : t \geq 0\}$ is a Poisson process with rate $\lambda p_j = p_j$. Let X_j denote the time when the first event occurs in the jth process, and $X = \max_{1 \leq j \leq m}$ be the time at which we complete a collection. Hence,

$$\mathbb{P}(X < t) = \mathbb{P}(X_j < t, \forall j = 1, \dots, m) = \prod_{j=1}^{m}(1 - e^{-p_j t}) \tag{2.4.4}$$

The expected value of X is thus given by

$$\mathbb{E}[X] = \int_0^\infty \mathbb{P}(X > t)dt = \int_0^\infty 1 - \prod_{j=1}^{m}(1 - e^{-p_j t})dt \tag{2.4.5}$$

To see the linkage between X and N, we let T_i be the interarrival time between the occurrence of the $i-1$th event and the ith event of the Poisson process N. Then $T_i \sim \exp(1)$. It is evident that given N coupons collected,

$$\mathbb{E}[X|N] = \mathbb{E}(\sum_{i=1}^{N} T_i|N) = N\mathbb{E}[T_i] = N \tag{2.4.6}$$

Therefore, $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}(X|N)] = \mathbb{E}[N]$.

Example 2.12 (**The Coupon Collecting Problem Continued).** Suppose now we are interested in the number of types that only appear once in the entire collection we gathered by the time we complete the coupon set. Let I_i be the indicator variable that

$$I_i = \begin{cases} 1, & \text{only a single coupon of type } i \text{ in the final collection} \\ 0, & \text{otherwise} \end{cases} \tag{2.4.7}$$

Hence, $I_i = 1$ means that for all $j \neq i$, the last time we collect a type j coupon precedes the second time that a coupon i will come up. That is, let S_i denote the time a second coupon of type i occurs, then $S_i \sim gamma(2, p_i)$ and

$$\begin{aligned} \mathbb{P}(I_i = 1) &= \mathbb{P}(X_j < S_i, \forall j \neq i) \\ &= \int_{t=0}^{\infty} \mathbb{P}(X_j < t, \forall j \neq i) p_i^2 t e^{-p_i t} dt \\ &= \int_{t=0}^{\infty} \prod_{j \neq i}(1 - e^{-p_j t}) p_i^2 t e^{-p_i t} dt \end{aligned} \tag{2.4.8}$$

Therefore,

$$\begin{aligned} \mathbb{E}[\sum_{i=1}^{m} I_i] &= \mathbb{E}[\sum_{i=1}^{m} \mathbb{P}(I_i)] \\ &= \sum_{i=1}^{m} \int_{t=0}^{\infty} \prod_{j \neq i}(1 - e^{-p_j t}) p_i^2 t e^{-p_i t} dt \\ &= \int_{t=0}^{\infty} \prod_{j=1}^{m}(1 - e^{-p_j t}) \sum_{i=1}^{m} p_i^2 t \frac{e^{-p_i t}}{1 - e^{-p_i t}} dt \end{aligned} \tag{2.4.9}$$

Let $N_1 = \{N_1(t) : t \geq 0\}$ and $N_2 = \{N_2(t) : t \geq 0\}$ be two independent Poisson process with respective rates λ_1 and λ_2. Let $S_n^{(1)}$ be the time at which n events occurred in the first process and $S_m^{(2)}$ the time at which m events occurred in the second process. From Equation 1.4.2, we can easily see that $\mathbb{P}(S_1^{(1)} < S_1^{(2)}) = \lambda_1/(\lambda_1+\lambda_2)$. Using the memoryless property, we have

$$\begin{aligned}\mathbb{P}(S_2^{(1)} < S_1^{(2)}) &= \mathbb{P}(S_2^{(1)} < S_1^{(2)}|S_1^{(1)} < S_1^{(2)})\mathbb{P}(S_1^{(1)} < S_1^{(2)}) \\ &= \mathbb{P}(S_1^{(1)} < S_1^{(2)})\mathbb{P}(S_1^{(1)} < S_1^{(2)}) \\ &= \left[\frac{\lambda_1}{\lambda_1+\lambda_2}\right]^2\end{aligned} \tag{2.4.10}$$

In fact, this means that each event can be viewed as occurring in a Poisson process $N = \{N(t) : t \geq 0\}$ with probability $p = \lambda_1/(\lambda_1 + \lambda_2)$ to be type I and probability $1 - p$ to be type II. Therefore, for general n and m, $\mathbb{P}(S_n^{(1)} < S_m^{(2)})$ is the probability that at least n type I events occurs out of the first $n + m - 1$ events. That is,

$$\mathbb{P}(S_n^{(1)} < S_m^{(2)}) = \sum_{k=n}^{n+m-1} \binom{n+m-1}{k} \Big(\frac{\lambda_1}{\lambda_1+\lambda_2}\Big)^k \Big(\frac{\lambda_2}{\lambda_1+\lambda_2}\Big)^{n+m-1-k} \tag{2.4.11}$$

2.5 Conditional Distribution of the Arrival Times ❖

Definition 2.13. Let $Y_1, Y_2, \ldots, Y_n$ be n random variables. $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ are the **order statistics** corresponding to $Y_1, Y_2, \ldots, Y_n$ if $Y_{(k)}, k = 1, 2, \ldots, n$ is the kth smallest value among $Y_1, Y_2, \ldots, Y_n$.

If $Y_1, Y_2, \ldots, Y_n$ are identically distributed independent random variables, then the joint density of the order statistics $Y_{(1)}, Y_{(2)}, \ldots, Y_{(n)}$ is given by

$$f(y_1, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i), \quad y_1 < y_2 < \ldots < y_n \tag{2.5.1}$$

If $Y_i, i = 1, \ldots, n$ is uniformly distributed in a time interval of length t, then the joint density is

$$f(y_1, y_2, \ldots, y_n) = n! \prod_{i=1}^{n} f(y_i) = \frac{n!}{t^n} \tag{2.5.2}$$

Theorem 2.14. *Given that $N(t) = n$, the n 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 $(0, t)$.*

Proof.

$$\begin{aligned}f(s_1, s_2, \ldots, s_n|N(t) = n) &= \frac{f(s_1, s_2, \ldots, s_n, N(t) = n)}{\mathbb{P}(N(t) = n)} \\ &= \frac{f(s_1, s_2, \ldots, s_n)}{\mathbb{P}(N(t) = n)} \\ &= \frac{\lambda^n e^{-\lambda s_1} e^{-\lambda(s_2-s_1)} \ldots e^{-\lambda(s_n-s_{n-1})} e^{-\lambda(t-s_n)}}{\frac{(\lambda t)^n}{n!} e^{-\lambda t}} \\ &= \frac{n!}{t^n}\end{aligned} \tag{2.5.3}$$

□

Proposition 2.15. *If $N_i(t), i = 1, \ldots, k$ represents the number of type i events occurring by time t, then $N_i(t), i = 1, \ldots, k$ are independent Poisson random variables having means*

$$\mathbb{E}[N_i(t)] = \lambda \int_0^t P_i(s) ds \tag{2.5.4}$$

where $P_i(s)$ is the probability of an event being type i if the event occurs at time s.

Proof. Consider the joint distribution that each type of event occurred for $n_i, n_i \geq 0$ times,

$$\mathbb{P}[N_i(t) = n_i, i = 1, \dots, k] = \mathbb{P}[N_i(t) = n_i, i = 1, \dots, k | N(t) = \sum_{i=1}^{k} n_i] \mathbb{P}[N(t) = \sum_{i=1}^{k} n_i] \tag{2.5.5}$$

Let P_i be the probability that an event occurring in the time interval $[0, t]$ is of type i. Due to the uniformness from Theorem 2.14, each event is equally likely to occur at any time in $s \in [0, t]$. We have

$$P_i = \frac{1}{t} \int_0^t P_i(s) ds \tag{2.5.6}$$

Again out of the independence of the stationary poisson process, we have the probability

$$\mathbb{P}[N_i(t) = n_i | N(t) = \sum_{i=1}^{k} n_i] = \frac{(\sum_{i=1}^{k} n_i)!}{n_1! \dots n_k!} P_1^{n_1} \dots P_k^{n_k} \tag{2.5.7}$$

is in effect a multinomial joint probability of the occurrence of k different types of events. Overall, we obtain

$$\begin{aligned} \mathbb{P}[N_i(t) = n_i, i = 1, \dots, k] &= \mathbb{P}[N_i(t) = n_i, i = 1, \dots, k | N(t) = \sum_{i=1}^{k} n_i] \mathbb{P}[N(t) = \sum_{i=1}^{k} n_i] \\ &= \frac{(\sum_{i=1}^{k} n_i)!}{n_1! \dots n_k!} P_1^{n_1} \dots P_k^{n_k} \frac{(\lambda t)^{\sum_{i=1}^{k} n_i}}{(\sum_{i=1}^{k} n_i)!} e^{-\lambda t} \\ &= \prod_{i=1}^{k} e^{-\lambda t P_i} \frac{(\lambda t P_i)^{n_i}}{n_i!} \end{aligned} \tag{2.5.8}$$

□

Example 2.16 (A one lane road with no overtaking). Suppose there is a single lane road with distance L and cars enters this road following a Poisson process with rate λ. Furthermore, we suppose that each car has a velocity $V_i, i = 1, 2, \dots$ that is independent and identically distributed. We further suppose that the distribution of V_i is independent from the counting process of cars entering the road. Note that one car with V_i needs to slow down to V_{i-1} if the preceding car is still on this road and $V_{i-1} < V_i$. To obtain the distribution of $R(t)$, the number of cars on this road at time t, we first denote $T_i = L/V_i$. That is, T_i equals the travel time of a car it is the only vehicle along the road. Let $G(x) = \mathbb{P}(T_i \leq x)$, we have

$$G(x) = \mathbb{P}(T_i \leq x) = \mathbb{P}(L/V_i \leq x) = \mathbb{P}(V_i \geq L/x) \tag{2.5.9}$$

We say that an event occurs when a car enters the road. Further, we say that an event is type I if it occurs at time $s \leq t$ and $L/V_i > t - s$. That is, the car is still travelling on the road at time t even if there is no slower car in front of it. Then a type I event occurring at time s has the probability

$$P(s) = \begin{cases} 1 - G(t - s) = \bar{G}(t - s), & \text{if } s \leq t \\ 0, & \text{if } s > t \end{cases} \tag{2.5.10}$$

Let $N_1(y)$ denote the number of type I event by time y, then by Proposition 2.15, $N_1(y)$ is a poisson process

$$\mathbb{E}[N_1(y)] = \lambda \int_0^y \bar{G}(t - s) ds, \quad y \leq t \tag{2.5.11}$$

Note that $R(t) = 0$ if and only if no type I events occur before time t. That is,

$$\mathbb{P}[R(t) = 0] = \mathbb{P}[N_1(t) = 0] = \exp[-\lambda \int_0^t \bar{G}(t - s) ds] = \exp[-\lambda \int_0^t \bar{G}(u) du] \tag{2.5.12}$$

To obtain $\mathbb{P}[R(t) = n], n > 0$, we need to condition on when the first type I event occurs. Since if the first car is still travelling, then the $n-1$ cars following it must still be on the road. Let X be the time when the first type I event occurs. Therefore,

$$F_X(y) = \mathbb{P}(X < y) = \mathbb{P}[N_1(y) > 0] = 1 - \exp[-\lambda \int_0^y \bar{G}(t-s)ds], \quad y \le t \tag{2.5.13}$$

Taking the derivative,

$$f_X(y) = \lambda G(t-y)\exp[-\lambda \int_0^y \bar{G}(t-s)ds], \quad y \le t \tag{2.5.14}$$

Therefore, conditioning on that the first type I event occurs at time y, $R(t)$ is just the first type I event plus a poisson process of cars entering the road. That is,

$$\mathbb{P}(R(t) = n | X = y) = e^{-\lambda(t-y)} \frac{[\lambda(t-y)]^{n-1}}{(n-1)!}, \quad \text{if } y \le t \tag{2.5.15}$$

Therefore,

$$\begin{aligned} \mathbb{P}(R(t) = n) &= \int_0^t \mathbb{P}(R(t) = n | X = y) f_X(y) dy \\ &= \int_0^t e^{-\lambda(t-y)} \frac{[\lambda(t-y)]^{n-1}}{(n-1)!} \lambda G(t-y) \exp[-\lambda \int_0^y \bar{G}(t-s)ds] dy \end{aligned} \tag{2.5.16}$$

3. Nonstationary Poisson Process

❖

Definition 3.1. The counting process $N = \{N(t), t \ge 0\}$ is said to be a **nonhomogeneous** (or **nonstationary**) **Poisson process** with intensity function $\lambda(t), t \ge 0$, if

1. $N(0) = 0$.
2. $$\mathbb{P}(N(t+h) = m+n | N(t) = n) = \begin{cases} \lambda(t)h + o(h), & \text{if } m = 1; \\ o(h), & \text{if } m > 1; \\ 1 - \lambda(t)h + o(h), & \text{if } m = 0. \end{cases} \tag{3.0.1}$$
3. It possesses independent increments.

The function

$$m(t) = \int_0^t \lambda(y) dy \tag{3.0.2}$$

is called the **mean value function**.

Theorem 3.2. *If $N = \{N(t), t \ge 0\}$ is a nonstationary Poisson process with intensity function $\lambda(t), t \ge 0$, then $N(t+s) - N(s)$ is a Poisson random variable with mean $m(t+s) - m(s) = \int_s^{t+s} \lambda(y)dy$.*

Proof. Mimicking the proof of Theorem 2.6, we have

$$p_j(0) = \begin{cases} 0 & \text{if } j \neq 0, \\ 1 & \text{if } j = 0. \end{cases} \tag{3.0.3}$$

We define the generating function

$$G(s,t) = \mathbb{E}[s^{N(t)}] = \sum_{j=0}^{\infty} p_j(t) s^j \tag{3.0.4}$$

Then,

$$\begin{aligned} G(s,t+h) &= \mathbb{E}[s^{N(t+h)}] \\ &= G(s,t)\mathbb{E}[s^{N(t+h)-N(t)}] \\ &= G(s,t)(1-\lambda(t)h+\lambda(t)hs+o(h)) \end{aligned} \tag{3.0.5}$$

Therefore,

$$\begin{aligned} G' &= \lim_{h\downarrow 0}\frac{G(s,t+h)-G(s,t)}{h} \\ &= \frac{G\lambda(t)h(s-1)+o(h)}{h} \\ &= G\lambda(t)(s-1) \end{aligned} \tag{3.0.6}$$

We move G to the left hand side and integrate on both sides

$$\int_0^t \frac{G'}{G}dt = \int_0^t \lambda(t)(s-1)dt \tag{3.0.7}$$

which gives us

$$\log(G(s,t))|_0^t = (s-1)\int_0^t \lambda(t)dt \tag{3.0.8}$$

Hence, we must have

$$G(s,t) = e^{(s-1)\int_0^t \lambda(t)dt} = e^{-(s-1)\int_0^t \lambda(t)dt}\sum_{j=0}^{\infty}\frac{((s-1)\int_0^t \lambda(t)dt)^j}{j!}s^j \tag{3.0.9}$$

which concludes the proof. □

4. Compound Poisson Process

❖

Definition 4.1. Let $X_1, X_2, \dots, X_n$ be independent random variables with the same distribution F with mean μ and variance σ^2. We say that the random variable $S = \sum_{i=1}^n X_i$ a **compound random variable**.

Definition 4.2. A stochastic process $X = \{X(t), t \geq 0\}$ is said to be a **compound Poisson process** if it can be represented as

$$X(t) = \sum_{i=1}^{N(t)} Y_i, t \geq 0 \tag{4.0.1}$$

where $N = \{N(t), t \geq 0\}$ is a Poisson process and $Y = \{Y_i, i \geq 1\}$ is a family of independent and identically distributed random variables that is also independent of $N = \{N(t), t \geq 0\}$.

References

[1] Grimmett, Geoffrey., and Stirzaker, David. Probability and Random Processes / Geoffrey R. Grimmett and David R. Stirzaker. 3rd ed., Oxford University Press, 2001.

[2] Resnick, Sidney I. Adventures in Stochastic Processes / Sidney Resnick. BirkhÃ¤user, 1992.

[3] Ross, Sheldon M. Introduction to Probability Models Sheldon M. Ross. 9th ed., Academic Press, 2007.