Lecture 7: Exponential distribution and Poisson process

1 Exponential Distribution

Definition 1.1. A random variable T is said to be *exponentially distributed* with rate λ, denoted as $T = \text{exponential}(\lambda)$, if the probability density function is

$$f_T(t) = \begin{cases} \lambda e^{-\lambda t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{1}$$

The cumulative distribution function of an exponentially distributed random variable is

$$F_T(t) = P[T \leq t] = \int_{-\infty}^{t} f_T(x)\, dx = 1 - e^{-\lambda t}. \tag{2}$$

Equivalently,

$$P[T > t] = e^{-\lambda t}. \tag{3}$$

Exponentially distributed random variables have the so-called memoryless property.

Definition 1.2. A random variable T is said to be *memoryless* if $P[T > t_1 + t_2 \mid T > t_1] = P[T > t_2]$.

Being memoryless means that the probability that some event happens after a certain time interval does not depend on what has happened before. To make it easier to understand, consider a bus station where buses arrive with exponentially distributed interarrival times with identical rates λ. If you have been waiting for a bus in vain for t_1 time, the probability that you will have to wait t_2 time for the next bus is the same as if you had just come to the station. In other words, you can forget about what has happened in the past.

As a matter of fact, exponential distribution is the only continuous probability distribution that has memoryless property. We state this as a theorem without proving it.

Theorem 1.3. *A continuous random variable X is memoryless if and only if it has an exponential distribution.*

Exponential distribution is widely used in the modelling and analysis of a wide variety of real world problems. For example, interarrival times of telephone calls at a local telephone switching office, or the service time of a bank counter, can often be represented exactly or approximately using the exponential distribution.

The expectation and variance of T can be derived by *moment generating function*

$$\Phi(x) = E[e^{xT}] = E[\sum_{n=0}^{\infty} \frac{(xT)^n}{n!}] = \sum_{n=0}^{\infty} E[T^n] \frac{x^n}{n!}$$

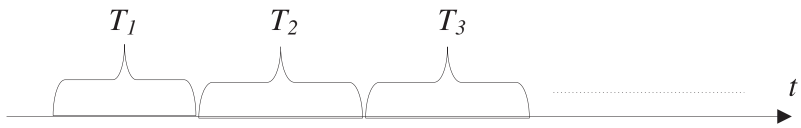

Figure 1: Independent waiting time/arrival interval

We also know that

$$
\begin{aligned}
E[e^{xT}] &= \int_{-\infty}^{\infty} e^{xt} f_T(t)\, dt \\
&= \frac{\lambda}{\lambda - x} \\
&= \frac{1}{1 - \frac{x}{\lambda}} \\
&= \sum_{n=0}^{\infty} (\frac{x}{\lambda})^n \\
&= \sum_{n=0}^{\infty} \frac{n!}{\lambda^n} \frac{x^n}{n!}
\end{aligned}
$$

Compare the coefficients before $\frac{x^n}{n!}$, we then have

$$E[T^n] = \frac{n!}{\lambda^n} \tag{4}$$

Hence,

$$E[T] = \frac{1}{\lambda} \tag{5}$$

$$Var[T] = E[T^2] - (E[T])^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \tag{6}$$

Remark 1.4. There are two forms of moment generating functions, namely, the normal generating function, which has the form of $\sum_{n=0}^{\infty} a_n x^n$, and the exponential generating function, which has the form of $\sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$.

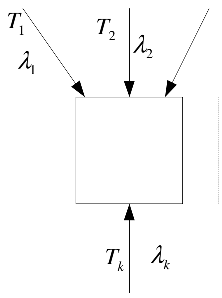

Figure 2: The exponential race problem

Example 1.5 (Exponential Race Problem). Several lines carrying incoming packet streams are connected to a router. Figure 2 illustrates the process. The interarrival time (first arrival time) of those packet streams are independent exponentially distributed random variables with rates $\lambda_1, \lambda_2, \cdots, \lambda_k$. $T_1, T_2, \cdots, T_k$ denote the first arrival time on line 1,2, $\cdots$,k.

The probability that the first packet arriving at the router is from line 1 is

$$\begin{aligned}
\text{Prob}\big[\text{first arrival is from line } 1\big] &= \text{Prob}\big[T_1 = \min\{T_1, \cdots, T_k\}\big] \\
&= \int_0^\infty P\big[T_2 > t\big] \cdots P\big[T_k > t\big] f_{T_1}(t)\, dt \\
&= \int_0^\infty e^{-(\lambda_1 + \cdots + \lambda_k)t} \lambda_1 e^{-\lambda_1 t}\, dt \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2 + \cdots + \lambda_k}.
\end{aligned}$$

When deriving the above formula, we use the fact that $\int_0^\infty e^{-at}\, dt = \frac{1}{a}$.

The first arrival time Z is given my $Z = \min\{T_1, \cdots, T_k\}$. The cumulative distribution function

$$\begin{aligned}
F_Z(t) = P[Z \leq t] &= 1 - P[Z > t] \\
&= 1 - \prod_1^k P[T_i > t] \\
&= 1 - e^{-(\lambda_1 + \cdots + \lambda_k)t}.
\end{aligned}$$

Hence, Z is also exponentially distributed with rate $(\lambda_1 + \cdots + \lambda_k)$.

Similarly, the last arrival time, i.e., $W = \max\{T^1, \cdots, T^k\}$, has a distribution function

$$F_W(t) = P[W \leq t] = \prod_1^k (1 - e^{-\lambda_i t}).$$

Unlike the first arrival time Z, the last arrival time W is not exponentially distributed. Intuitively, it is because that max function renders W dependent on what happened in the past, thus not being memoryless.

Example 1.6 (Gamma Distribution). The n^{th} arrival interval S_n is the time when the n^{th} packet arrives:

$$S_n = T_1 + \cdots + T_n$$

.

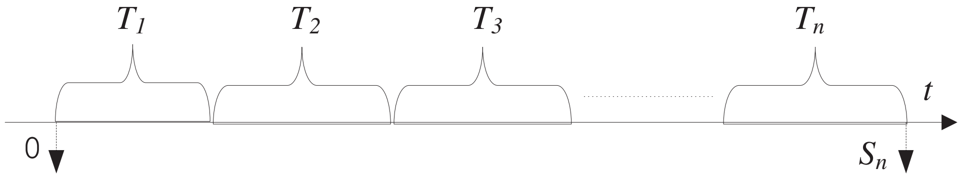

Figure 3: The n^{th} arrival interval

The cumulative distribution function

$$F_{S_n}(t) = P[S_n \leq t] = \int_0^t (P[S_n - T_1 \leq t - x]) f_{T_1}(x)\, dx$$

.

From the above equation, we can derive the probability density function of S_n

$$f_{S_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}. \tag{7}$$

S_n is called a *Gamma random variable*, denoted as $\mathrm{Gamma}(n, \lambda)$.

2 Poisson Process

Again we look at the random process where points arrive randomly with independent identically exponentially distributed arrival intervals.

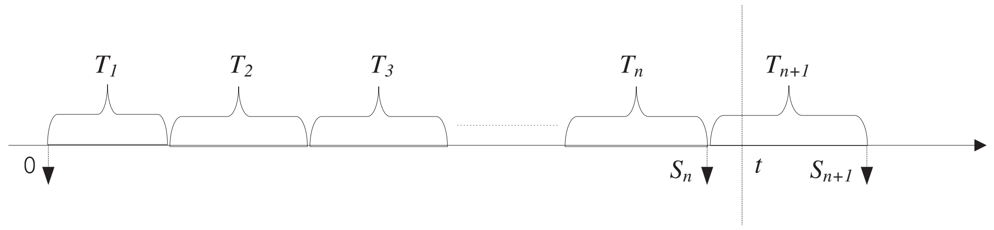

Figure 4: Poisson Process

In Figure 4, arrivial intervals $T_1, T_2, \cdots, T_n, T_{n+1}$ are all i.i.d. random variables that have an exponential distribution with rate λ. S_n is the arrival time of the n^{th} point.

Definition 2.1. A random process $\{N(t), t \geq 0\}$ defined as $N(t) = \max\{n \colon S_n \leq t\}$ is called a Poisson process with rate λ, written as Poisson(λt).

The distribution function is

$$\begin{aligned} P[N(t) = n] &= P[S_n \leq t, S_{n+1} > t] \\ &= \int_0^t (P[T_{n+1} > t - x]) f_{s_n}(x)\, dx \\ &= \cdots \\ &= e^{-\lambda t} \cdot \frac{(\lambda t)^n}{n!}. \end{aligned}$$

We may rewrite the above formula by substituting μ for λt, and rename $N(t)$ as X. X is called a *Poisson random variable with parameter* $\mu > 0$ if

$$P[x = n] = e^{-\mu} \frac{\mu^n}{n!}. \tag{8}$$

It is denoted as $X = \mathrm{Poisson}(\mu)$.

It is well know that the expectation of a Poisson random variable with parameter μ is

$$E[X] = \mu,$$

and the variance is

$$Var[x] = \mu.$$

Proposition 2.2. *If* X = *Poisson*(μ), Y = *Poisson*(λ), *and* X, Y *are indpendent, then* $X + Y$ = *Poisson*$(\mu + \lambda)$.

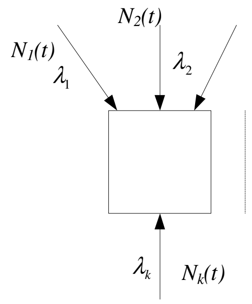

Figure 5: Merging (Superposition) of Poisson Processes

Similarly, let $N(t) = \sum N_i(t)$. $N_(t)$ is called the *merging*, or *superposition* of Poisson processes, as illustrated in Figure 5.

Proposition 2.3. *The merging, or superposition of independent Poission processes $N_1(t), N_2(t), \cdots, N_k(t)$ with rates $\lambda_1, \lambda_2, \cdots, \lambda_k$ is a new Poisson process $N(t)$ with rate $\lambda = \sum_{i=1}^{k} \lambda_i$.*

Proof.

$$\begin{aligned} P[N(t) = n] &= P[\sum_{i=1}^{k} N_i(t) = n] \\ &= P[\sum_{i=1}^{k} X_i = n] \\ &= e^{-(\lambda_1+\cdots+\lambda_k)t} \frac{[(\lambda_1 + \cdots + \lambda_k)t]^n}{n!} \end{aligned}$$

Hence $N(t) = \text{Poisson}(\lambda t), \text{ where} \lambda = \sum_{i=1}^{k} \lambda_i$. □

There are several equivalent definitions for Poisson process. Below is another commonly used definition.

Definition 2.4. A stochastic process $N(t)$ is a Poission process if it satisfies the following conditions:

1. $N(0) = 0$;
2. $N(t + s) - N(s) = \text{Poisson}(\lambda t)$;
3. $N(t)$ has *independent increments*, i.e.,$\forall t_0 < t_1 < \cdots < t_n$, $N(t_1) - N(t_0)$,$N(t_2) - N(t_1)$, $\cdots$,$N(t_n) - N(t_{n-1})$ are independent;

Indpendent Poisson processes may be merged (superpositioned) to form a new Poisson process whose rate is the sum of the rates of those merged (superpositioned) Poisson processes. If we reverse the direction of flow of packets in Figure 5, we call the problem *splitting*, or *thinning* of Poisson process, as depicted in Figure 6. The packets in the main line are distributed probabilistically into several lines. $N(t)$, which is defined as the number of packets up till time t, is a Poisson process with rate λ. The probability that a packet is fed into the i^{th} line is p_i, and $\sum_{i=1}^{k} p_i = 1$. If a packet is fed into line i, it is called a packet of *type i*. $N_i(t)$ is defined as the number of packets of *type i* up till time t.

Proposition 2.5. *$N_i(t)$ is a Poisson process with rate λp_i.*

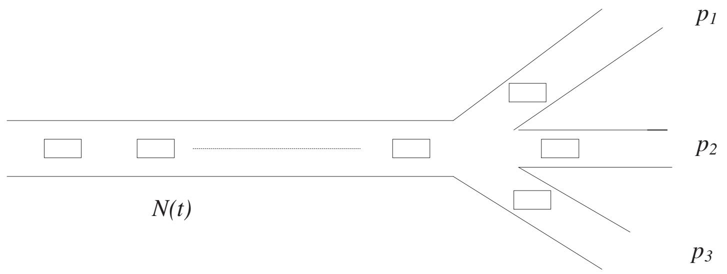

Figure 6: Splitting (Thinning) of a Poisson Process

Proof. The joint probability function of $N_1(t), N_2(t), \cdots, N_k(t)$ is

$$\begin{aligned} P[N_1(t)=n_1,\cdots,N_k(t)=n_k] &= \sum_{n=0}^{\infty} P[N_1(t)=n_1,\cdots,N_k(t)=n_k \mid N(t)=n]P[N(t)=n] \\ &= P[N_1(t)=n_1,\cdots,N_k(t)=n_k \mid N(t)=n_0]P[N(t)=n_0] \\ &= \binom{n_0}{n_1,\cdots,n_k} p_1^{n_1}\cdots p_k^{n_k} e^{-\lambda t}\frac{(\lambda t)^{n_0}}{n_0!} \\ &= \prod_{j=1}^{k} e^{-\lambda p_j t}\frac{(\lambda p_j t)^{n_j}}{n_j!}. \end{aligned}$$

We use n_0 to denote $n_1+\cdots+n_k$ in the above derivation.
The probability function of $N_1(t)$ may be written as

$$\begin{aligned} P[N_1(t)=n] &= \sum_{n_2,\cdots,n_k} P[N_1(t)=n, N_2(t)=n_2,\cdots,N_k(t)=n_k] \\ &= \sum_{n_2,\cdots,n_k} e^{-\lambda p_1 t}\frac{(\lambda p_1 t)^n}{n!}\prod_{j=2}^{k} e^{-\lambda p_j t}\frac{(\lambda p_j t)^{n_j}}{n_j!} \\ &= e^{-\lambda p_1 t}\frac{(\lambda p_1 t)^n}{n!}\prod_{j=2}^{k}\sum_{n_j=0}^{\infty} e^{-\lambda p_j t}\frac{(\lambda p_j t)^{n_j}}{n_j!}. \end{aligned}$$

Note the fact that

$$\sum_{i,j,k} a_i b_j c_k = (\sum a_i)(\sum b_j)(\sum c_k).$$

That's why we can exchange $\prod$ and $\sum$ in the above formula. Also note that

$$\sum_{n_j=0}^{\infty} e^{-\lambda p_j t}\frac{(\lambda p_j t)^{n_j}}{n_j!} = 1,$$

the probability function of $N_1(t)$ is

$$P[N_1(t)=n] = e^{-\lambda p_1 t}\frac{(\lambda p_1 t)^n}{n!}.$$

Hence $N_1(t)$ is a Poisson process with rate λ_1. $\square$