Lecture notes 7
The Exponential Distribution and the Poisson Process

Definition:
A continuous random variable X is said to have an exponential distribution with parameter $\lambda,\ \lambda > 0$, if its probability density function is given by

$$f(x) = \lambda e^{-\lambda x},\ x \geq 0.$$

Its cumulative distribution function is given by

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

Note that $E(X) = \frac{1}{\lambda}$ and $Var(X) = \frac{1}{\lambda^2}$.

Properties of the Exponential Distribution:

Definition:
A random variable X is said to be without memory (i.e. memoryless), if

$$P(X > s + t \mid X > t) = P(X > s)$$

for all $s,\ t \geq 0$.

Proof:

$$P(X > s + t \mid X > t) = \frac{P(X > s + t)}{P(X > t)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda t}} = P(X > s)$$

Remark: The exponential distributed random variables are memoryless.

Example:

Suppose that the amount of time one spends in a bank is exponentially distributed with mean ten minutes, that is, $\lambda = \frac{1}{10}$.

1. What is the probability that a customer will spend more than fifteen minutes in the bank?
2. What is the probability that a customer will spend more than fifteen minutes in the bank given that he is still in the bank after ten minutes?

Solution:

If X represents the amount of time that the customer spends in the bank, then the first probability is just

$$P(X > 15) = e^{-15\lambda} = e^{-3/2} \approx 0.220$$

The second question asks for the probability that a customer who has spent ten minutes in the bank will have to spend at least five more minutes. However, since the exponential distribution does not "remember" that the customer has already spent ten minutes in the bank, this must equal the probability that an entering customer spends at least five minutes in the bank. That is, the desired probability is just

$$P(X > 5) = e^{-5\lambda} = e^{-1/2} \approx 0.604$$

Example:

Suppose that the amount of time that a lightbulb works before burning itself out is exponentially distributed with mean ten hours. Suppose that a person enters a room in which a lightbulb is burning. If this person desires to work for five hours, then what is the probability that he will be able to complete his work without the bulb is burning out? What can be said about this probability when the distribution is not exponential?

Solution:

Since the bulb is burning when the person enters the room it follows, by the memoryless property of the exponential, that its remaining lifetime is exponential with mean ten. Hence the desired probability is

$$P(\text{remaining lifetime} > 5) = 1 - F(5) = e^{-5\lambda} = e^{-1/2}$$

However, if the lifetime distribution F is not exponential, then the relevant probability is

$$P(\text{lifetime} > t + 5 \mid \text{lifetime} > t) = \frac{1 - F(t+5)}{1 - F(t)}$$

where t is the amount of time that the bulb had been in use prior to the person entering the room. That is, if the distribution is not exponential then additional information is needed (namely, t) before the desired probability can be calculated. In fact, it is for this reason, namely, that the distribution of the remaining lifetime is independent of the amount of time that object has already survived, that the assumption of an exponential distribution is so often made.

The Poisson Process

1. Counting Processes

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 have occurred up to time t. Some examples of counting processes are the following:

(a) If we let $N(t)$ equal the number of persons who have entered a particular store at or prior to time t, then $\{N(t), t \geq 0\}$ is a counting process in which an event corresponds to a person entering the store. Note that if we had let $N(t)$ equal the number of persons in the store at time t, then $\{N(t), t \geq 0\}$ would *not* be a counting process. (Why not?)

(b) If we say that an event occurs whenever a child is born, then $\{N(t), t \geq 0\}$ is a counting process when $N(t)$ equals the total number of people who were born by time t. (Does $N(t)$ include persons who have died by time t? Explain why it must.)

(c) If $N(t)$ equals the number of goals that a given soccer player has scored by time t, then $\{N(t), t \geq 0\}$ is a counting process. An event of this process will occur whenever the soccer player scores a goal.

From this definition we see that for a counting process $N(t)$ must satisfy:

i. $N(t) \geq 0$
ii. $N(t)$ is integer valued.
iii. If $s < t$, then $N(s) \leq N(t)$.
iv. For $s < t$, $N(t) - N(s)$ equals the number of events that have occurred in the interval (s, t).

A counting process is said to possess independent increments if the numbers of events which occur in disjoint time intervals are independent. For example, this means that the number of events which have occurred by time 10 [that is, $N(10)$] must be independent of the number of events occurring between times 10 and 15 [that is, $N(15) - N(10)$].

• The assumption of independent increments might be reasonable for example (a), but it probably would be unreasonable for example (b). The reason for this is that if in example (b) $N(t)$ is very large, then it is probable that there are many people alive at time t; this would lead us to believe that the number of new births between t and $t + s$ would also tend to be large [that is, it does not seem reasonable that $N(t)$ is independent of $N(t + s) - N(t)$, and so $N(t), t \geq 0$ would not have independent increments in example (b)].
• The assumption of independent increment in example (c) would be justified if we believe that the soccer player's chances of scoring a goal today does not depend on "how he's been going". It would not be justified if we believed in "hot streaks" or "slumps".

A counting process is said to possess stationary increments if the distribution of the number of events which 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 (t_1+s, t_2+s)(that is, $N(t_2+s)-N(t_1+s)$) has the same distribution as the same distribution as the number of events in the interval (t_1, t_2)(that is, $N(t_2)-N(t_1)$) for all $t_1 < t_2$, and $s > 0$.

• The assumption of stationary increments would only be reasonable in example (a) if there were no times of day at which people were more likely to enter the store. Thus, for instance, if there was a rush hour (say, between 12 P.M. and 1 P.M.) each day, then the stationarity assumption would not be justified.
• If we believed that the earth's population is basically constant (a belief not held at present by most scientists), then the assumption of stationary increments might be reasonable in example (b).
• Stationary increments do not seem to be a reasonable assumption in example (c) since, for one thing, most people would agree that the soccer player would probably score more goals while in the age bracket 25-30 than he would while in the age bracket 35-40. *It may, however, be reasonable over a smaller time horizon, such as one year.*

2. Definition of the Poisson Process

– *Patience Pays.*

The Poisson distribution is often referred to as *law of rare events.* Specifically, it is the macro-law (distribution) of micro-rare events. In general, rareness in micro scales can aggregate to become common in a macro scale. In short, *macro-common is aggregated micro-rareness.* For example, air crash is rare in the micro scale of per flight or per hour, but not so in the macro-scale of all flights around the world or of decades in time. The so-called "law of rare events" can be accurately depicted with the binomial approximation, which is essential in understanding Poisson random variables and Poisson processes.

Example 1*. (Traffic Accidents) Suppose that

(1). The chance of one traffic accident on the Tolo Highway on any one day is very small, so small that more than one traffic accidents one day is ignorable);

(2). Over different days, the occurrence of traffic accidents are independent;

(3). On average, there are 3 traffic accidents on the Tolo Highway per year.

Then, X, the random number of traffic accidents over one year, follows $Bin(365, p)$ with $p = 3/365$. And X follows approximately the Poisson distribution with mean 3, i.e.,

$$X \sim \mathcal{P}(3), \qquad \text{approximately.}$$

Example 2*. (CHELSEA VS. LIVERPOOL) Chelsea plays against Liverpool this Saturday. Suppose

(1). the chance for Chelsea to score a goal within any given one minute is very small; (Scoring 2 or more goals within one minute is so rare that it's ignorable.)

(2). Over different minutes, scoring of goals are independent.

Then, the number of goals of Chelsea in the match is a Poisson random variable, a random variable following a Poisson distribution.

One of the most important counting process is the Poisson process which is defined as follows:

Definition (1):
The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, if

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

$$P[N(t+s) - N(s) = n] = e^{-\lambda t} \frac{(\lambda t)^n}{n!}, \quad n = 0, 1, \ldots$$

Note that if follows from condition (iii) that a Poisson process has stationary increments and also that

$$E[N(t)] = \lambda t$$

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

To determine if an arbitrary counting process is actually a Poisson process, we must show that conditions (i), (ii), and (iii) are satisfied. Condition (i), which simply states that the counting of event begins at time $t = 0$, and condition (ii) can usually be directly verified from our knowledge of the process. However, it is not at all clear how we would determine that condition (iii) is satisfied, and for this reason an equivalent definition of a Poisson process would be useful.

Definition: The function $f(\cdot)$ is said to be $o(h)$ if

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

Examples:

(a) The function $f(x) = x^2$ is $o(h)$ since

$$\lim_{h \to 0} \frac{f(h)}{h} = \lim_{h \to 0} \frac{h^2}{h} = \lim_{h \to 0} h = 0$$

(b) The function $f(x) = x$ is not $o(h)$ since

$$\lim_{h\to 0} \frac{f(h)}{h} = \lim_{h\to 0} \frac{h}{h} = \lim_{h\to 0} 1 = 1 \neq 0$$

(c) If $f(\cdot)$ is $o(h)$ and $g(\cdot)$ is $o(h)$, then so is $f(\cdot) + g(\cdot)$. This follows since

$$\lim_{h\to 0} \frac{f(h) + g(h)}{h} = \lim_{h\to 0} \frac{f(h)}{h} + \lim_{h\to 0} \frac{g(h)}{h} = 0 + 0 = 0$$

(d) If $f(\cdot)$ is $o(h)$, then so is $g(\cdot) = cf(\cdot)$. This follows since

$$\lim_{h\to 0} \frac{cf(h)}{h} = c \lim_{h\to 0} \frac{f(h)}{h} = c \cdot 0 = 0$$

(e) From (iii) and (iv) it follows that any finite linear combination of functions, each of which is $o(h)$, is $o(h)$.

Remark: In order for the function $f(\cdot)$ to be $o(h)$, it is necessary that $f(h)/h$ go to zero as h goes to zero. But if h goes to zero, the only way for $f(h)/h$ to go to zero is for $f(h)$ to go to zero faster than h does. That is, for h to be small, $f(h)$ must be small compared with h.

Definition (2):
The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process having rate λ, $\lambda > 0$, if

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

Theorem: The above two definitions of a Poisson process are equivalent.

Example 2* (Continued). (CHELSEA VS. LIVERPOOL) Suppose, in addition to the assumptions (1) and (2) in Example 2, the chance, denoted by p, for Chelsea to score a goal within any given one minute is equal. Let $X(t)$ be the number of the Chelsea goals from the beginning of the match till time t, with the unit of the time being minute. Then, $\{X(t) : t \in [0, 90]\}$ is *approximately* a Poisson process with rate p, the mean number of goals per minute, the time unit. (A match has 90 minutes.)

The Poisson process can be viewed as tallying micro-rare events over time. Therefore it is widely used to model number of events which happen randomly along time. Here λ measures the intensity of the events happening. It measures the average number of events per time unit. The larger (smaller) the λ, the more (less) likely the events. The most essential ingredient of the Poisson process is that *over different non-overlapping time intervals, event occurrences are entirely independent.* More general version of Poisson processes allows the intensity vary over time, which is referred to as non-homogeneous Poisson process.

3. Interarrival (Sojourn) and Waiting Time Distributions

Consider a Poisson process, and let us denote the time of the first event by T_1. Further, for $n > 1$, let T_n denote the elapsed time between the $(n-1)$st and the nth event. The sequence $\{T_n,\ n = 1, 2, \ldots\}$ is called the *sequence of interarrival times.*

For instance, if $T_1 = 5$ and $T_2 = 10$, then the first event of the Poisson process would have occurred at time 5 and the second at time 15. We shall now determine the distribution of the T_n. To do so, we first note that the event $\{T_1 > t\}$ takes place if and only if no events of the Poisson process occur in the interval $[0, t]$ and thus,

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

Hence, T_1 has an exponential distribution with mean $1/\lambda$. Now,

$$P(T_2 > t) = E[P(T_2 > t \mid T_1)]$$

Reason: Let E_1 be the event $\{T_2 > t\}$. Let $X = 1$ if E_1 occurs, and $X = 0$ if E_1 does not occur. Then

$$E(X) = P(E_1) = P(T_2 > t)$$

and

$$E(X|T_1) = P(E_1|T_1) = P(T_2 > t \mid T_1).$$

Therefore,

$$P(T_2 > t) = E(X) = E[E(X|T_1)] = E[P(T_2 > t \mid T_1)].$$

However,

$$\begin{aligned} P(T_2 > t|T_1 = s) &= P(\ 0 \ \text{ events in } (s, s+t]\ |T_1 = s) \\ &= P(\ 0 \ \text{ events in } (s, s+t]) \\ &= e^{-\lambda t} \end{aligned} \tag{1}$$

where the last two equations followed from independent and stationary increments. Therefore, from Equation (1) we conclude that T_2 is also an exponential random variable with mean $1/\lambda$ and, furthermore, that T_2 is independent of T_1. Repeating the same argument yields the following.

Proposition 1
$\{T_n,\ n = 1, 2, \ldots\}$ are independent identically distributed exponential random variables having mean $1/\lambda$.

Remarks
The proposition should not surprise us. The assumption of stationary and independent increments is basically equivalent to asserting that, at any point in time, the process

probabilistically restarts itself. That is, the process from any point on is independent of all that has previously occurred (by independent increments), and also has the same distribution as the original process (by stationary increments). In other words, the process has *no memory*, and hence exponential interarrivial times are to be expected.

Another quantity of interest is W_n, the arrival time of the nth event, also called the waiting time until the nth event. It is easily seen that

$$W_n = \sum_{i=1}^{n} T_i, \qquad n \geq 1$$

and hence from Proposition 1 and the results in the review of probability that W_n has a gamma distribution with parameters n and λ. That is, the probability density function of W_n is given by

$$f_{W_n} = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!}, \qquad t \geq 0.$$

This can be derived by noting that the nth event will occur prior to or at time t if and only if the number of events occurring by time t is at least n. That is

$$N(t) \geq n \iff W_n \leq t.$$

Hence,

$$\begin{aligned} F_{W_n}(t) = P(W_n \leq t) &= P(N(t) \geq n) \\ &= \sum_{k=n}^{\infty} \frac{(\lambda t)^k e^{-\lambda t}}{k!} = 1 - \sum_{k=0}^{n-1} \frac{(\lambda t)^k e^{-\lambda t}}{k!}. \end{aligned}$$

Then, the density of W_n is

$$\begin{aligned} f(t) = \frac{d}{dt} P(W_n \leq t) &= -\sum_{k=0}^{n-1} \frac{d}{dt} \left(\frac{(\lambda t)^k e^{-\lambda t}}{k!}\right) \\ = \quad & -\sum_{k=1}^{n-1} \frac{k \lambda^k t^{k-1} e^{-\lambda t}}{k!} - \sum_{k=0}^{n-1} \frac{(\lambda t)^k (-\lambda) e^{-\lambda t}}{k!} = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}. \end{aligned}$$

Example

Suppose that people immigrate into a territory at a Poisson rate $\lambda = 1$ per day.

(a) What is the expected time until the tenth immigrant arrives?

(b) What is the probability that the elapsed time between the tenth and the eleventh arrival exceeds two days?

Solution

(a) $E(W_{10}) = E(\sum_{i=1}^{10} T_i) = 10/\lambda =$10 days.

(b) $P(T_{11} > 2) = e^{-2\lambda} = e^{-2} \approx 0.133$.