Lecture 5

Exponential Distribution and Poisson Process

- Exponential distribution and its properties memoryless
- Poisson process
- Interarrival time and waiting time
- Compound Poisson process

5.1 The Exponential Distribution

5.1.1 Definition

- Exponential distribution is used to model the time between events. It is characterized by the rate (average number of events per unit time).
- A continuous random variable X is said to have an *exponential distribution* with *rate* λ, $\lambda > 0$, if its pdf is given by

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

- Notation: $X \sim \text{Exp}(\lambda)$
- The cdf is given by

$$P\{X \leq x\} = F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

$$P\{X > x\} = 1 - F(x) = e^{-\lambda x}$$

- Expected value:

$$E[X] = \int_0^\infty x \cdot \lambda e^{-\lambda x} dx = \frac{1}{\lambda}.$$

- Variance:

$$\begin{aligned} E[X^2] &= \int_0^\infty x^2 \cdot \lambda e^{-\lambda x} dx = \frac{2}{\lambda^2} \\ \mathrm{Var}(X) &= E[X^2] - E[X]^2 = \frac{1}{\lambda^2}. \end{aligned}$$

5.1.2 Properties of the Exponential Distribution

- A random variable X is said to be *memoryless* if

$$P\{X > t + s | X > t\} = P\{X > s\} \qquad \text{for all } s, t \geq 0.$$

- The above condition is equivalent to

$$P\{X > t + s\} = P\{X > t\}P\{X > s\}$$

- The exponential distribution $P\{X > s\} = e^{-\lambda s}$ satisfies the condition, and is therefore memoryless.

- In fact, exponential distribution is the only Continous distribution that is memoryless, i.e., $g(x) = e^{-\lambda x}$ is the only function that satisfies

$$g(s + t) = g(s)g(t) \qquad \text{for all } s, t \geq 0.$$

Example 5.1 Suppose that the amount of time X one spends in a bank is exponentially distributed with mean ten minutes, that is, $\lambda = \frac{1}{10}$ (number of persons per minute). What is the probability that a customer will spend more than fifteen minutes in the bank? What is the probability that a customer will spend more than fifteen minutes in the bank given that she is still in the bank after ten minutes?

$$f(x) = \begin{cases} 10e^{-10x} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

$E(X) = E_Y(E_X(X|Y))$

Proposition 5.2 (Conditional Variance Formula)

- $\mathrm{Var}(X) = \underset{Y}{E}[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y]).$

$x_1, x_2, \cdots x_n$

$\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} = s^2$

Proof:

$\mathrm{Var}(X|Y) \qquad E(X-EX)^2$

$$
\begin{aligned}
E[\mathrm{Var}(X|Y)] &= E\left[E[X^2|Y] - E[X|Y]^2\right] \\
&= E[X^2] - E\left[E[X|Y]^2\right] \\
\underbrace{\mathrm{Var}(E[X|Y])}_{Z} &= E\left[\underbrace{E[X|Y]}_{Z}{}^2\right] - (E[\underbrace{E[X|Y]}_{Z}])^2 \\
&= E\left[E[X|Y]^2\right] - E[X]^2.
\end{aligned}
$$

total exp. formula

$x_1, x_2, \cdots X_N$

$\dfrac{\sum_{i=1}^{N} (x_i - \mu)^2}{N} = \sigma^2$

$= E(X^2|Y) - (E(X|Y))^2$

$= E(E(X|Y))$

$= EX$

Hence,

$$E[\mathrm{Var}(X|Y)] + \mathrm{Var}(E[X|Y]) = E[X^2] - E[X]^2.$$

□

Example 5.3 A store must decide how much of a certain commodity to order so as to meet next month's demand, where that demand is assumed to have an exponential distribution with rate λ. If the commodity costs the store c per pound, and can be sold at a price of $s > c$ per pound, how much should be ordered so as to maximize the store's expected profit?

Solution: Let X equal the demand, so that $X \sim \text{Exp}(\lambda)$. If the store orders the amount t, then the profit, call it P, is given by

$$P = s \cdot \min\{X, t\} - ct.$$

profit = revenue - cost

① demand $x >$ ordered t

$ts - t \cdot c$

② demand $x <$ ordert t

$x \cdot s - t \cdot c$

Hence,

$f(t)$ $f'(t)=0 \Rightarrow t=t^*$

$$E[P] = s \cdot E[\min\{X, t\}] - ct.$$

To calculate $E[\min\{X, t\}]$, we condition on whether $X > t$:

Finally, we have

$$\begin{aligned} E[P] &= s \cdot E[\min\{X, t\}] - ct \\ &= s\left[te^{-\lambda t} + \frac{1}{\lambda} - \left(t + \frac{1}{\lambda}\right)e^{-\lambda t}\right] - ct \\ &= \frac{s}{\lambda} - \frac{s}{\lambda}e^{-\lambda t} - ct. \end{aligned}$$

To maximize the expected profit

$$f(t) = \frac{s}{\lambda} - \frac{s}{\lambda}e^{-\lambda t} - ct,$$

we solve $f'(t) = 0$:

$$f'(t) = se^{-\lambda t} - c = 0,$$

so that

$$t = -\frac{1}{\lambda}\ln\left(\frac{c}{s}\right).$$

□

$X \sim Exp(\lambda)$

Example 5.4 The dollar amount of damage involved in an automobile accident is an exponential random variable with mean 1000, i.e. $\lambda = 1/1000$. Of this, the insurance company only pays the amount exceeding (the deductible amount of) 400. Find ① the expected value and the standard deviation ② of the amount the insurance company pays per accident.

$E(Y) = E(Y|X>400)\cdot P(X>400) + E(Y|X\le 400)\cdot P(X\le 400)$

$= E(X-400|X>400)\cdot P\{X<400\} + 0.$

Solution: Let X be the dollar amount of damage resulting from an accident. Let

$= (E(X|X>400) - 400)\cdot P\{X>400\}.$

$$Y = \begin{cases} X - 400, & \text{if } X > 400 \\ 0, & \text{if } X \le 400. \end{cases}$$

$= \{400 + 1000 - 400\}\cdot e^{-\frac{1}{1000}\cdot 400}$

≈ 670.32

be the amount paid. Since $X \sim \text{Exp}(1/1000)$, it is memoryless. Thus,

$$P\{Y > y|X > 400\} = P\{X > 400 + y|X > 400\} = P\{X > y\},$$

so that $Y \sim \text{Exp}(1/1000)$ when $X > 400$ is given. On the other hand, if $X \le 400$ is given, then $Y \equiv 0$.

Consider the random variable I defined by

$$I = \begin{cases} 1, & \text{if } X > 400 \\ 0, & \text{if } X \le 400. \end{cases}$$

Such I is therefore a Bernoulli random variable with success probability

$$P\{I = 1\} = P\{X > 400\} = e^{-0.4}.$$

Next, consider the random variable $Z = E[Y|I]$, its value is

$$Z = \begin{cases} 1000, & \text{if } I = 1 \\ 0, & \text{if } I = 0. \end{cases}$$

Therefore, the expected payment by the insurance company per accident is

$$\begin{aligned} E[Y] &= E[E[Y|I]] = E[Z] \\ &= 1000 \times e^{-0.4} + 0 \times (1 - e^{-0.4}) \approx 670.32. \end{aligned}$$

To calculate $\mathrm{Var}(Y)$, we apply the Conditional Variance formula,

$$\mathrm{Var}(Y) = E[\mathrm{Var}(Y|I)] + \mathrm{Var}(E[Y|I]).$$

Consider the random variable $W = \mathrm{Var}[Y|I]$ defined by

$$W = \begin{cases} 1000^2, & \text{if } I = 1 \\ 0, & \text{if } I = 0. \end{cases}$$

Var(Y) = Var(X-400) = Var(X)

$= \frac{1}{\lambda^2}$

The variance $\mathrm{Var}(Y)$ is thus equal to

$$\begin{aligned}
\mathrm{Var}(Y) &= E[W] + \mathrm{Var}(Z) \\
&= E[W] + E[Z^2] - E[Z]^2 \\
&= 1000^2 \times e^{-0.4} + 0 \times (1 - e^{-0.4}) \\
&\quad + 1000^2 \times e^{-0.4} + 0^2 \times (1 - e^{-0.4}) - \left[1000 \times e^{-0.4} + 0 \times (1 - e^{-0.4})\right]^2 \\
&= 10^6 \times e^{-0.4} + 10^6 \times e^{-0.4} - 10^6 \times e^{-0.8} = 10^6 \times (2 \times e^{-0.4} - e^{-0.8}).
\end{aligned}$$

The standard deviation of Y is thus

$$\sqrt{\mathrm{Var}(Y)} \approx 944.09.$$

□

5.2 The Poisson Process

5.2.1 Counting Process

- A *counting process* $\{N(t)|t \geq 0\}$ gives the total number of "events" that occur by time t.

- It is a continuous-time stochastic process that satisfies:

  (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 occur during $(s, t]$.

Example 5.5

1. Let $N(t)$ equal the number of persons who enter a particular store at or prior to time t. Then $\{N(t)|t \geq 0\}$ is a counting process. The total of customers in the store at time t is, however, not a counting process.

2. If we say that an event occurs whenever a child is born, then the process $\{N(t)|t \geq 0\}$ is a counting process when $N(t)$ equals the total number of people who were born by time t.

3. If $N(t)$ equals the number of goals that a given soccer player scores by time t, then $\{N(t)|t \geq 0\}$ is a counting process.

$\square$

- For $s < t$, the value $N(t) - N(s)$ is called an *increment* during $(s, t]$.
- A counting process is said to possess *independent increments* if the numbers of events that occur in disjoint time intervals are independent.
- E.g. Having independent increments would mean that the number of customers arriving a store between 2p.m. and 3p.m. is independent of that between 3p.m. and 4p.m. But that between 2p.m. and 3p.m. is dependent on that between 1p.m. and 3p.m.

- 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.
- In other words, the process has stationary increments if $N(s+t) - N(s)$ has the same distribution for all s.
- E.g. If the number of customers arriving a store has stationary increments, then the number of customers arrived within one hour has a fixed distribution, regardless of the time of the day.

5.2.2 Poisson Process

- We adopt a simple definition of a Poisson process.
- A counting process $\{N(t)|t \geq 0\}$ is said to be a *Poisson process* with rate $\lambda > 0$ if the following additional conditions hold:
  1. $N(0) = 0$;
  2. $\{N(t)|t \geq 0\}$ has independent increments;
  3. The increment $N(s+t) - N(s)$ is Poisson distributed with mean λt, i.e.

$$P\{N(s+t) - N(s) = n\} = e^{-\lambda t} \cdot \frac{(\lambda t)^n}{n!} \qquad n \geq 0.$$

- Condition 3 implies that the increments are stationary because the distribution is independent of s.
- When $s = 0$, we have $P\{N(t) = n\} = e^{-\lambda t} \cdot \frac{(\lambda t)^n}{n!}$.

Example 5.6 Suppose that people immigrate into a territory at a Poisson rate $\lambda = 2$ per day.

(a) How many immigrants are expected in a period of 10 days?

Solution: $E[N(10)] = 2 \cdot 10 = 20$ $\square$

(b) What is the probability of no immigrants in 5 days?

Solution: $P\{N(5) = 0\} = e^{-2\times 5} = 0.0000454$ $\square$

(c) What is the probability of more than one immigrant in 5 days?

5.2.3 Interarrival and Waiting Time Distributions

- Consider a Poisson process. Denote the time of the first event by T_1. Let T_n denote the elapsed time between the $(n-1)$st and the nth event. The sequence $\{T_n | n \geq 1\}$ is called the sequence of *interarrival times*.

- E.g. If $T_1 = 5$ and $T_2 = 10$, then the first event occurs at time 5 and the second at time 15.

- Distribution of T_1:

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

  Therefore $T_1 \sim \text{Exp}(\lambda)$.

- Distribution of T_2:

$$\begin{aligned}
P\{T_2 > t\} &= \int_0^\infty P\{T_2 > t | T_1 = s\} f_{T_1}(s) ds \\
&= \int_0^\infty P\{N(s+t) - N(s) = 0 | T_1 = s\} f_{T_1}(s) ds \\
&= \int_0^\infty P\{N(s+t) - N(s) = 0\} f_{T_1}(s) ds \\
&= \int_0^\infty e^{-\lambda t} f_{T_1}(s) ds \\
&= e^{-\lambda t} \int_0^\infty f_{T_1}(s) ds \\
&= e^{-\lambda t}.
\end{aligned}$$

Therefore $T_2 \sim \text{Exp}(\lambda)$.

- In general, all interarrival times are i.i.d. exponentially distributed random variables with mean $1/\lambda$.

- The *waiting time* until the nth event is

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

- Distribution of S_n:

$$F_{S_n}(t) = P\{S_n \leq t\} \;=\; P\{N(t) \geq n\} = \sum_{i=n}^{\infty} e^{-\lambda t} \cdot \frac{(\lambda t)^i}{i!}.$$

Thus,

$$\begin{aligned} f_{S_n}(t) = F'_{S_n}(t) \;&=\; \sum_{i=n}^{\infty} (-\lambda) e^{-\lambda t} \cdot \frac{(\lambda t)^i}{i!} + \sum_{i=n}^{\infty} e^{-\lambda t} \cdot \lambda \frac{(\lambda t)^{i-1}}{(i-1)!} \\ &=\; \lambda e^{-\lambda t} \cdot \frac{(\lambda t)^{n-1}}{(n-1)!}. \end{aligned}$$

- We say that S_n is *gamma* distributed with parameters n and λ.

Example 5.7 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?

Solution: $E[S_{10}] = \frac{10}{\lambda} = 10$ (days) $\square$

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

5.2.4 Conditional Distribution of the Arrival Times

Example 5.8 Suppose that one event occurs by time t. Determine the distribution of the waiting time T_1.

Solution: For $s \le t$,

$$\begin{aligned} P\{T_1 < s | N(t) = 1\} &= \frac{P\{T_1 < s,\ N(t) = 1\}}{P\{N(t) = 1\}} \\ &= \frac{P\{N(s) = 1,\ N(t) - N(s) = 0\}}{P\{N(t) = 1\}} \\ &= \frac{P\{N(s) = 1\} \cdot P\{N(t) - N(s) = 0\}}{P\{N(t) = 1\}} \\ &= \frac{P\{N(s) = 1\} \cdot P\{N(t - s) = 0\}}{P\{N(t) = 1\}} \\ &= \frac{\lambda s e^{-\lambda s}/1! \cdot e^{-\lambda(t-s)}/0!}{\lambda t e^{-\lambda t}/1!} = \frac{s}{t}. \end{aligned}$$

Hence, it is uniformly distributed over $(0, t)$. □

Example 5.9 Determine $P\{N(s) = 1, N(t) = 2\}$ for $s < t$.

independent increasement

$P\{N(s)=1, N(t)=2\} = P\{N(s)=1\} \cdot P\{N(t)-N(s)=1\}$

$P\{N(s)=1\} = \frac{(\lambda s)^1 \cdot e^{-\lambda s}}{1!} = \lambda s e^{-\lambda s}$

$P\{N(t)-N(s)=1\} = \frac{[\lambda(t-s)]^1 e^{-\lambda(t-s)}}{1!} = \lambda(t-s)e^{-\lambda(t-s)}$

multiplying, $P\{N(s)=1, N(t)=2\} = \lambda^2 s(t-s)e^{-\lambda t}$

Example 5.10 Note that

$$
\begin{aligned}
P\{N(5) = 2, N(10) - N(5) = 1\} &= P\{N(5) = 2\}P\{N(5) = 1\} \\
P\{N(5) = 2, N(10) - N(5) = 1\} &\neq P\{N(5) = 2, N(5) = 1\}.
\end{aligned}
$$

5.2.5 Compound Poisson Process

- A stochastic process $\{X(t)|t \geq 0\}$ is said to be a *compound Poisson process* if it can be represented as

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

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

Example 5.11

1. If $Y_i \equiv 1$, then $X(t) = N(t)$, and so we have the usual Poisson process.

2. Suppose that buses arrive at a sporting event in accordance with a Poisson process, and suppose that the numbers of fans in each bus are assumed to be i.i.d. Then $\{X(t)|t \geq 0\}$ is a compound Poisson process where $X(t)$ denotes the total number of fans who have arrived by time t. Here, Y_i represents the number of fans in the ith bus.

3. Suppose that customers leave a supermarket in accordance with a Poisson process. If the Y_i, the amount spent by the ith customer, $i = 1, 2, \ldots,$ are i.i.d., then $\{X(t)|t \geq 0\}$ is a compound Poisson process when $X(t)$ denotes the total amount of money spent by time t. $\square$

- Since N is independent of $\{Y_i | i \geq 1\}$, we have (cf. Example 2.5)

$$
\begin{aligned}
E[X(t)] = E[E[X(t)|N]] &= \sum_{n=0}^{\infty} E[X(t)|N = n] \cdot P\{N = n\} \\
&= \sum_{n=0}^{\infty} E\left[\textstyle\sum_{i=1}^{n} Y_i \middle| N = n\right] \cdot P\{N = n\} \\
&= \sum_{n=0}^{\infty} E\left[\textstyle\sum_{i=1}^{n} Y_i\right] \cdot P\{N = n\} \\
&= \sum_{n=0}^{\infty} n \cdot E[Y] \cdot P\{N = n\} \\
&= E[Y] \cdot \sum_{n=0}^{\infty} n \cdot P\{N = n\} \\
&= E[Y] \cdot E[N] = \lambda t \cdot E[Y].
\end{aligned}
$$

- Likewise, the variance can be obtained as $\mathrm{Var}(X(t)) = \lambda t \cdot E[Y^2]$.

Example 5.12 Suppose that families migrate to an area at a Poisson rate $\lambda = 2$ per week. If the number of people in each family is independent and takes on the values 1, 2, 3, 4 with respective probabilities $\frac{1}{6}$, $\frac{1}{3}$, $\frac{1}{3}$, $\frac{1}{6}$, then what is the expected value and variance of the number of individuals migrating to this area during a fixed five-week period?