Chapter 4

Continuous Random Variables

A random variable can be discrete, continuous, or a mix of both. Discrete random variables are characterized through the probability mass functions, i.e., the ideal histograms. However, the same argument does not hold for continuous random variables because the width of each histogram's bin is now infinitesimal. In this chapter we will generalize PMF to a new concept called probability density function, and derive analogous properties.

4.1 Probability Density Function

To understand how a continuous random variable can be characterized, we start by generalizing the probability mass function.

Let X be a discrete random variable. In principle, we should consider the states of X as a *countable sequence* because X is discrete. However, for the time being let us consider X as continuous. Let x_0 be one of its states, and suppose that there is an interval $[a, b]$ such that for any $x \in [a, b]$,

$$p_X(x) = \begin{cases} 0, & \text{if} \quad x \neq x_0, \\ p_0, & \text{if} \quad x = x_0. \end{cases}$$

Then, according to the definition of PMF, the probability of $a \leq X \leq b$ is defined as

$$\mathbb{P}[a \leq X \leq b] \overset{(a)}{=} \mathbb{P}[X = x_0] = p_0 \overset{(b)}{=} \int_a^b \underbrace{p_0 \; \delta(x - x_0)}_{f_X(x)} dx. \tag{4.1}$$

In this equation, (a) holds because there is no other non-zero probabilities in the interval $[a, b]$ besides x_0. The integration in (b) follows from the definition of a delta function, which is that $\delta(x - x_0) = \infty$ if $x = x_0$, $\delta(x - x_0) = 0$ if $x \neq x_0$, and

$$\int_{-\infty}^{\infty} \delta(x - x_0) dx = 1.$$

That is, if we integrate the delta function then we will obtain one. If we further define $f_X(x) \stackrel{\text{def}}{=} p_0\, \delta(x - x_0)$, then the right hand side of Equation (4.1) provides a *density* characterization of the random variable X: When we integrate the density function $f_X(x)$, we will obtain the probability. Figure 4.1 shows an illustration.

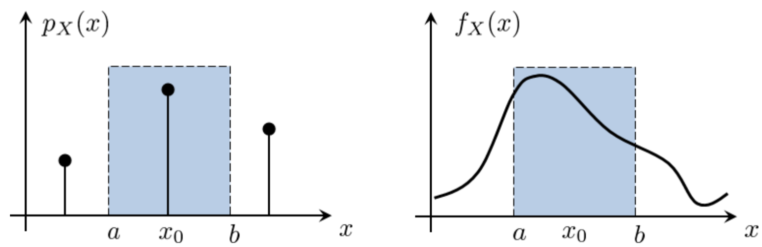

Figure 4.1: PMF is a train of impulses, whereas PDF is (usually) a smooth function.

Continuous random variables have a smooth density function as illustrated on the right hand side of Figure 4.1. The characterization, however, is the same as Equation (4.1). The following definition summarizes this.

Definition 1. *The* ***probability density function (PDF)*** *of a random variable X is a function which, when integrated over an interval $[a, b]$, yields the probability of obtaining $a \leq X \leq b$. We denote PDF of X as $f_X(x)$, and*

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f_X(x) dx. \tag{4.2}$$

Example 1. Let X be the phase angle of a voltage signal. Without any prior knowledge about X we may assume that X has an equal probability of any value between 0 to 2π. Find the PDF of X and compute $\mathbb{P}[0 \leq X \leq \pi/2]$.

Solution. Since X has an equal probability for any value between 0 to 2π, the PDF of X is

$$f_X(x) = \frac{1}{2\pi}, \qquad \text{for } 0 \leq x \leq 2\pi.$$

Therefore, the probability $\mathbb{P}[0 \leq X \leq \pi/2]$ can be computed as

$$\mathbb{P}\left[0 \leq X \leq \frac{\pi}{2}\right] = \int_0^{\pi/2} \frac{1}{2\pi} dx = \frac{1}{4}.$$

Remark: Note that when specifying the PDF of a continuous random variable, the range is very important, e.g., $0 \leq x \leq 2\pi$ in this example.

Example 2. Let X be a discrete random variable with PMF

$$p_X(k) = \frac{1}{2^k}, \qquad k = 1, 2, \ldots,$$

What is $f_X(x)$,the continuous representation of $p_X(k)$? Find the probability $\mathbb{P}[1 \le X \le 3]$.

Solution. The continuous representation of the PMF can be written in terms of a train of delta functions:

$$f_X(x) = \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right) \delta(x-k).$$

To see why this is the correct PDF, we can check a few x's:

$$\begin{aligned} f_X(1) &= \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right) \delta(1-k) = \frac{1}{2}, \\ f_X(2) &= \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right) \delta(2-k) = \frac{1}{4}, \end{aligned}$$

where in both examples $\delta(1-k) = 1$ only when $k = 1$ and $\delta(2-k) = 1$ only when $k = 2$. The probability $\mathbb{P}[1 \le X \le 3]$ can be computed as

$$\begin{aligned} \mathbb{P}[1 \le X \le 3] = \int_1^3 f_X(x)dx &= \int_1^3 \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right) \delta(x-k)dx \\ &= \sum_{k=1}^{\infty} \left(\frac{1}{2^k}\right) \int_1^3 \delta(x-k)dx = \sum_{k=1}^{3} \left(\frac{1}{2^k}\right) = \frac{7}{8}. \end{aligned}$$

Normalization Property

Same as the discrete random variable, the continuous random variable should have its PDF integrated to one.

Theorem 1. *A PDF $f_X(x)$ should satisfy*

$$\int_{-\infty}^{\infty} f_X(x)dx = 1. \tag{4.3}$$

Proof. The proof follows from the fact that

$$\int_{-\infty}^{\infty} f_X(x)dx = \mathbb{P}[-\infty \le X \le \infty],$$

which must be 1 as x is integrated over the entire real line. □

Example 3. Let $f_X(x) = c(1-x^2)$ for $-1 \le x \le 1$, and 0 otherwise. Find the constant c.

Solution. Since

$$\int_{-\infty}^{\infty} f_X(x)dx = \int_{-1}^{1} c(1-x^2)dx = \frac{4c}{3},$$

and because $\int_{-\infty}^{\infty} f_X(x)dx = 1$, we have $c = 3/4$.

What is $\mathbb{P}[X = x_0]$ if X is continuous?
The long answer will be discussed in the next section. The short answer is: If the PDF $f_X(x)$ (which is a function of x) is *continuous* at $x = x_0$, then $\mathbb{P}[X = x_0] = 0$. This can be seen from the definition of the PDF. As $a \to x_0$ and $b \to x_0$, the integration interval becomes 0. Thus, unless $f_X(x)$ is a delta function at x_0, which has been excluded because we assumed $f_X(x)$ is continuous at x_0, the integration result must be zero.

4.2 Cumulative Distribution Function

Definition 2. *The* ***cumulative distribution function (CDF)*** *of a continuous random variable X is*

$$F_X(x) \overset{def}{=} \mathbb{P}[X \le x] = \int_{-\infty}^{x} f_X(x')dx'. \tag{4.4}$$

If we compare this definition with the one for discrete random variable, we see that the CDF for a continuous random variable X simply replaces the summation of a discrete random variable by integration. Therefore, we should expect more of the properties to inherit from the discrete CDF.

Example. Let X be a continuous random variable with PDF

$$f_X(x) = \frac{1}{b-a}, \qquad a \le x \le b.$$

Find the CDF of X.

Solution. The CDF of X is given by

$$F_X(x) = \int_{-\infty}^{x} f_X(x')dx' = \int_{a}^{x} \frac{1}{b-a}dx' = \frac{x-a}{b-a},$$

for $a \le x \le b$. For $x < a$, $F_X(x) = 0$, and for $x > b$, $F_X(x) = 1$.

Properties of CDF

We now describe the properties of a CDF.

Proposition 1. *Let X be a random variable (either continuous or discrete), then the CDF of X has the following properties:*

(i) The CDF is a ***non-decreasing****.*

(ii) The ***maximum*** *of the CDF is when $x = \infty$: $F_X(+\infty) = 1$.*

(iii) The ***minimum*** *of the CDF is when $x = -\infty$: $F_X(-\infty) = 0$.*

Pictorially, these properties are illustrated from Figure 4.2: (i) follows from the fact that we are integrating a non-negative function $f_X(x)$. Thus the bigger the x the more areas we will integrate under the curve. (ii) and (iii) indicate the two extreme values of the CDF.

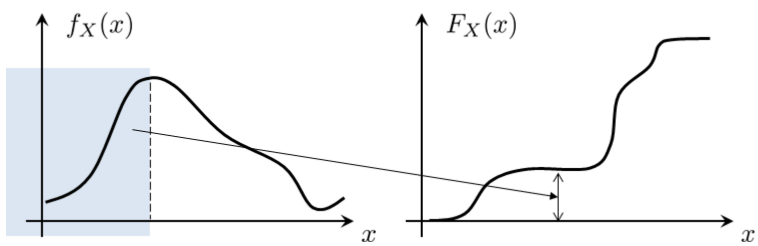

Figure 4.2: A CDF is the integration of the PDF. Thus, the height of a stem in the CDF corresponds to the area under the curve of the PDF.

The next property relates probability and CDF.

Proposition 2. *Let X be a continuous random variable. If the CDF F_X is continuous at any $a \leq x \leq b$, then*

$$\mathbb{P}[a \leq X \leq b] = F_X(b) - F_X(a). \tag{4.5}$$

Proof. The proof follows from the definition of the CDF, which states that

$$F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(x')dx' - \int_{-\infty}^{a} f_X(x')dx' = \int_{a}^{b} f_X(x')dx' = \mathbb{P}[a \leq X \leq b].$$

□

This result shows that for a continuous random variable X, $\mathbb{P}[X = x_0] = F_X(x_0) - F_X(x_0) = 0$. (That is, substitute $a = x_0$ and $b = x_0$ in the above equation.) Intuitively, this can be explained from the fact that the integration interval is infinitesimally small.

The next two propositions concern about the discontinuity of the CDF.

Definition 3. *The CDF $F_X(x)$ is said to be*

- ***Left-continuous*** *at $x = b$ if $F_X(b) = F_X(b^-) \overset{def}{=} \lim_{h\to 0} F_X(b-h)$;*
- ***Right-continuous*** *at $x = b$ if $F_X(b) = F_X(b^+) \overset{def}{=} \lim_{h\to 0} F_X(b+h)$;*
- ***Continuous*** *at $x = b$ if it is both right-continuous and left-continuous at $x = b$. In this case, we have*

$$\lim_{h\to 0} F_X(b-h) = \lim_{h\to 0} F_X(b+h) = F(b).$$

In this definition, the step size $h > 0$ is shrinking to zero. The point $b - h$ stays at the left of b, and $b + h$ stays at the right of b. The limits on the left hand side and the right hand side are not necessarily equal, unless the function $F_x(x)$ is continuous at $x = b$.

Proposition 3. *For any random variable X (discrete or continuous), $F_X(x)$ is* ***right-continuous****. That is,*

$$F_X(b) = F_X(b^+) \overset{def}{=} \lim_{h\to 0} F_X(b+h) \tag{4.6}$$

Right-continuous essentially means that if $F_X(x)$ is piecewise, then it must have a **solid left end and an empty right end**. If $F_X(x)$ is continuous at $x = b$, then the leftmost solid dot will overlap with the rightmost empty dot of the previous segment. Thus, there is no gap between the two. A pictorial illustration is shown in Figure 4.3.

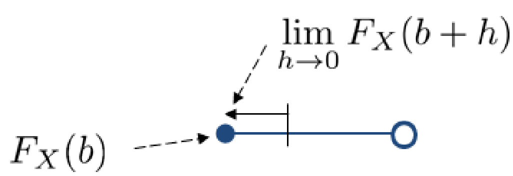

Figure 4.3: Illustration of right continuous.

Proposition 4. *For any random variable X (discrete or continuous), $\mathbb{P}[X = b]$ is*

$$\mathbb{P}[X = b] = \begin{cases} F_X(b) - F_X(b^-), & \textit{if } F_X \textit{ is discontinuous at } x = b \\ 0, & \textit{otherwise.} \end{cases} \tag{4.7}$$

This proposition concerns about the probability at discontinuity. It states that when $F_X(x)$ is discontinuous at $x = b$, then $\mathbb{P}[X = b]$ is the difference between $F_X(b)$ and the limit from the left. In other words, the height of the gap determines the probability at discontinuity. If $F_X(x)$ is continuous at $x = b$, then $F_X(b) = \lim_{h\to 0} F_X(b-h)$ and so $\mathbb{P}[X = b] = 0$.

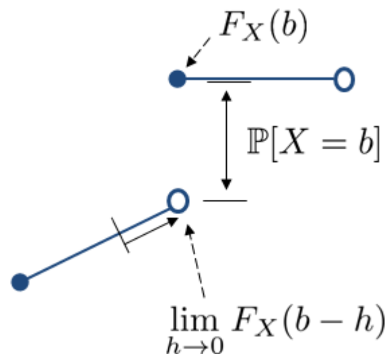

Figure 4.4: Illustration of Equation (4.7).

Converting between PDF and CDF

Thus far we have only seen how to obtain $F_X(x)$ from $f_X(x)$. In order to go in the reverse direction, we need to recall the fundamental theorem of calculus. Fundamental theorem of calculus states that if a function f is continuous, then (See Chapter 1)

$$f(x) = \frac{d}{dx}\int_a^x f(t)dt$$

for some constant a. Using this result for CDF and PDF, we have the following result:

Theorem 2. *The* ***probability density function*** *(PDF) is the derivative of the cumulative distribution function (CDF):*

$$f_X(x) = \frac{dF_X(x)}{dx} = \frac{d}{dx}\int_{-\infty}^x f_X(x')dx', \tag{4.8}$$

provided F_X is differentiable at x.

Remark: If F_X is not differentiable at x, then,

$$f_X(x) = \mathbb{P}[X = x] = F_X(x) - \lim_{h\to 0} F_X(x-h),$$

where the second equality follows from Equation (4.7).

Example 1. Consider a CDF

$$F_X(x) = \begin{cases} 0, & x < 0, \\ 1 - \frac{1}{4}e^{-2x}, & x \geq 0. \end{cases}$$

Find the PDF $f_X(x)$.

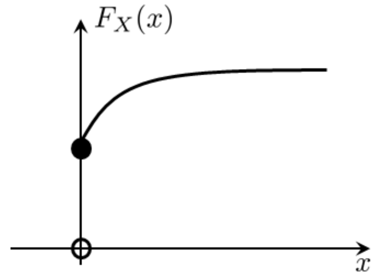

Figure 4.5: Example 1.

Solution. First, it is easy to show that $F_X(0) = \frac{3}{4}$. This corresponds to a discontinuity at $x = 0$, as shown in Figure 4.5. Because of the discontinuity, we need to consider three when determining $f_X(x)$:

$$f_X(x) = \begin{cases} \frac{dF_X(x)}{dx}, & x < 0, \\ \mathbb{P}[X = 0], & x = 0, \\ \frac{dF_X(x)}{dx}, & x > 0. \end{cases}$$

When $x < 0$, $F_X(x) = 0$. So, $\frac{dF_X(x)}{dx} = 0$.
When $x > 0$, $F_X(x) = 1 - \frac{1}{4}e^{-2x}$. So, $\frac{dF_X(x)}{dx} = \frac{1}{2}e^{-2x}$.
When $x = 0$, the probability $\mathbb{P}[X = 0]$ is determined according to Property 7 which is the height between the solid dot and the empty dot. This yields

$$\mathbb{P}[X = 0] = F_X(0) - \lim_{h \to 0} F_X(0 - h) = \frac{3}{4} - 0 = \frac{3}{4}.$$

Therefore, the overall PDF is

$$f_X(x) = \begin{cases} 0, & x < 0, \\ \frac{3}{4}, & x = 0, \\ \frac{1}{2}e^{-2x}, & x > 0. \end{cases}$$

4.3 Expectation, Variance and Moments

Definition 4. *The expectation of a continuous random variable X is*

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx. \tag{4.9}$$

If g is a function, then

$$\mathbb{E}[g(X)] = \int_{-\infty}^{\infty} g(x)\, f_X(x) dx. \tag{4.10}$$

Example 1. Let Θ be a continuous random variable with PDF $f_\Theta(\theta) = \frac{1}{2\pi}$. Let $Y = \cos(\omega t + \Theta)$. Find $\mathbb{E}[Y]$.

Solution. Referring to Equation (4.10), the function g is $g(\theta) = \cos(\omega t + \theta)$. Therefore, the expectation $\mathbb{E}[Y]$ is

$$\mathbb{E}[Y] = \int_0^{2\pi} \cos(\omega t + \theta) f_\Theta(\theta) d\theta = \frac{1}{2\pi} \int_0^{2\pi} \cos(\omega t + \theta) d\theta = 0,$$

where the last equality holds because integrating a sinusoid over one period is 0.

Example 2. Let $\mathbb{I}_\Omega(X)$ be an indicator function such that

$$\mathbb{I}_\Omega(X) = \begin{cases} 1, & \text{if} \quad X \in \Omega, \\ 0, & \text{if} \quad X \notin \Omega. \end{cases}$$

Find $\mathbb{E}[\mathbb{I}_\Omega(X)]$.

Solution. The expectation is

$$\mathbb{E}[\mathbb{I}_\Omega(X)] = \int_{-\infty}^{\infty} \mathbb{I}_\Omega(x) f_X(x) dx = \int_{x \in \Omega} f_X(x) dx = \mathbb{P}[X \in \Omega].$$

This result shows that the probability of an event $\{X \in \Omega\}$ can be equivalently represented in terms of expectation.

Existence of Expectation

Not all random variables has expectation. Only random variables that are absolutely integrable can have expectation.

Definition 5. *A random variable X is said to be* ***absolutely integrable*** *if*

$$\mathbb{E}[|X|] = \int_{-\infty}^{\infty} |x| f_X(x) dx < \infty. \tag{4.11}$$

Before we discuss an example, we should clarify that the limits of the integration in Equation (4.9) should be sent to infinity *separately*. That is,

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\, f_X(x) dx = \lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b x\, f_X(x) dx.$$

Note that

$$\lim_{a \to -\infty} \lim_{b \to \infty} \int_a^b x\, f_X(x) dx \neq \lim_{T \to \infty} \int_{-T}^{T} x\, f_X(x) dx.$$

The interpretation of the double limit is that we can decompose

$$\int_{-\infty}^{\infty} x\, f_X(x)dx = \lim_{a\to-\infty}\int_a^0 x\, f_X(x)dx + \lim_{b\to\infty}\int_0^b x\, f_X(x)dx,$$

and each integral has to be finite in order to ensure the expectation exists.

Example 3. Let X be a Cauchy random variable with PDF

$$f_X(x) = \frac{1}{\pi(1+x^2)}. \tag{4.12}$$

Show that the expectation of X does not exist.

Solution. From the definition of expectation, we know that

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} x\cdot\frac{1}{\pi(1+x^2)}dx = \frac{1}{\pi}\int_0^{\infty}\frac{x}{(1+x^2)}dx + \frac{1}{\pi}\int_{-\infty}^{0}\frac{x}{(1+x^2)}dx.$$

The first integral gives

$$\int_0^{\infty}\frac{x}{(1+x^2)}dx = \frac{1}{2}\log(1+x^2)\Big|_0^{\infty} = \infty,$$

and the second integral gives $-\infty$. Since neither integral is finite, the expectation is undefined. We can also check the absolutely integrable criteria:

$$\begin{aligned}\mathbb{E}[|X|] = \int_0^{\infty}|x|\cdot\frac{1}{\pi(1+x^2)}dx &\overset{(a)}{=} 2\int_0^{\infty}\frac{x}{\pi(1+x^2)}dx \ge 2\int_1^{\infty}\frac{x}{\pi(1+x^2)}dx \\ &\overset{(b)}{\ge} 2\int_1^{\infty}\frac{x}{\pi(x^2+x^2)}dx = \frac{1}{\pi}\log(x)\Big|_1^{\infty} = \infty,\end{aligned}$$

where in (a) we use the fact that the function being integrated is even, and in (b) we lower bound $\frac{1}{1+x^2} \ge \frac{1}{x^2+x^2}$ if $x > 1$.

Variance

Definition 6. *The variance of a continuous random variables X is*

$$\mathrm{Var}[X] = \mathbb{E}[(X-\mu_X)^2] = \int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx, \tag{4.13}$$

where $\mu_X \overset{def}{=} \mathbb{E}[X]$.

It is not difficult to show that the variance can also be expressed as

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mu_X^2,$$

because

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

4.4 Mean, Mode, Median

There are three statistical quantities we are usually interested in: Mean, mode and median. It is possible to obtain these three quantities from PDF and CDF.

Mean

The mean of a random variable is also the expectation of the random variable. Therefore, the mean can be computed via the PDF as follows.

Definition 7. *Let X be a random variable. The mean of X is*

$$\text{Mean} = \mathbb{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx. \tag{4.14}$$

Mean can also be computed from the CDF, shown in the following theorem.

Theorem 3. *The mean of a random variable X can be computed from the CDF as*

$$\text{Mean} = \mathbb{E}[X] = \int_0^{\infty} \left(1 - F_X(x')\right) dx' - \int_{-\infty}^{0} F_X(x') dx'. \tag{4.15}$$

Proof. For any random variable X, we can partition $X = X^+ - X^-$ where X^+ and X^- are the positive and negative parts, respectively. Consider the positive part first. For any $X > 0$, it holds that

$$\begin{aligned}
\int_0^{\infty} \left(1 - F_X(x')\right) dx' &= \int_0^{\infty} \left[1 - \mathbb{P}[X \le x']\right] dx' = \int_0^{\infty} \mathbb{P}[X > x'] dx' \\
&= \int_0^{\infty} \int_{x'}^{\infty} f_X(x) dx dx' = \int_0^{\infty} \int_0^{x} f_X(x) dx' dx \\
&= \int_0^{\infty} x f_X(x) dx = \mathbb{E}[X].
\end{aligned}$$

Now, consider the negative part. For any $X < 0$, it holds that

$$\begin{aligned}
\int_{-\infty}^{0} F_X(x') dx' &= \int_{-\infty}^{0} \mathbb{P}[X \le x'] dx' = \int_{-\infty}^{0} \int_{-\infty}^{x'} f_X(x) dx dx' \\
&= \int_{-\infty}^{0} \int_{x}^{0} f_X(x) dx' dx = \int_{-\infty}^{0} x f_X(x) dx = \mathbb{E}[X],
\end{aligned}$$

where in both cases we switched the order of the integration. □

Median

Definition 8. *The median of a random variable X is the point $-\infty < c < \infty$ such that*

$$\int_{-\infty}^{c} f_X(x)dx = \int_{c}^{\infty} f_X(x)dx. \tag{4.16}$$

Essentially, this means that the areas under the PDF on both sides of $x = c$ are equal.

The median can also be evaluated from the CDF as follows.

Theorem 4. *The median of a random variable X is the point c such that*

$$F_X(c) = \frac{1}{2}. \tag{4.17}$$

Proof. Since $F_X(x) = \int_{-\infty}^{x} f_X(x')dx'$, we have

$$F_X(c) = \int_{-\infty}^{c} f_X(x)dx = \int_{c}^{\infty} f_X(x)dx = 1 - F_X(c).$$

Rearranging the terms shows that $F_X(c) = \frac{1}{2}$. □

Mode

Definition 9. *The mode of a random variable X is the point c such that $f_X(x)$ attains the maximum:*

$$c = \underset{x}{\mathit{argmax}} \;\; f_X(x). \tag{4.18}$$

The mode of a random variable is not unique, e.g., a mixture of two identical Gaussians with different means has two modes.

Theorem 5. *The mode of a random variable X is the point c such that the slope of $F_X(x)$ is maximized.*

$$c = \underset{x}{\mathit{argmax}} \;\; F_X'(x). \tag{4.19}$$

Proof. Note that from the definition of CDF, we have

$$f_X(x) = \frac{d}{dx}\int_{-\infty}^{x} f_X(x')dx' = \frac{d}{dx}F_X(x) = F_X'(x).$$

Therefore,

$$c = \underset{x}{\operatorname{argmax}} \;\; f_X(x) = \underset{x}{\operatorname{argmax}} \;\; F_X'(x).$$

□

A pictorial illustration of mode and median is given in Figure 4.6.

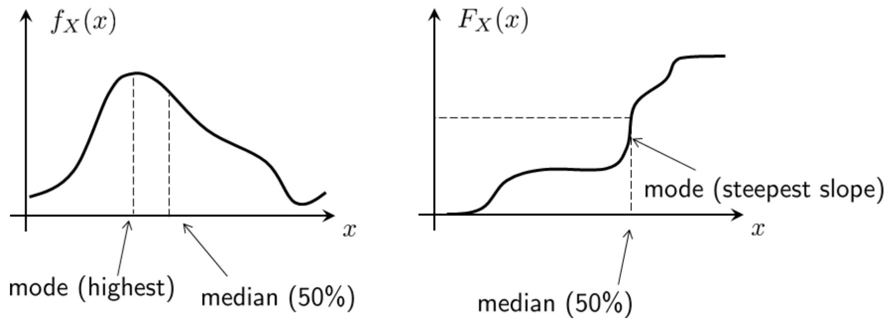

Figure 4.6: Mode and median in a PDF and a CDF.

4.5 Common Continuous Random Variables

Uniform Random Variable

Definition 10. *Let X be a continuous uniform random variable. The PDF of X is*

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a \le x \le b, \\ 0, & \textit{otherwise}, \end{cases} \tag{4.20}$$

where $[a, b]$ is the interval on which X is defined. We write

$$X \sim \text{Uniform}(a, b)$$

to say that X is drawn from a uniform distribution on an interval $[a, b]$.

Uniform distribution can also be defined for discrete random variables. In this case, the probability mass function is given by

$$p_X(k) = \frac{1}{b-a+1}, \quad k = a, a+1, \dots, b.$$

The presence of "1" in the denominator of the PMF is due to the fact that k runs from a to b, including the two end points.

The mean and variance of a uniform random variable is stated in the theorem below.

Theorem 6. *If $X \sim \text{Uniform}(a, b)$, then*

$$\mathbb{E}[X] = \frac{a+b}{2}, \quad \textit{and} \quad \text{Var}[X] = \frac{(b-a)^2}{12}.$$

Proof. The proof follows from the definition of expectation and variance:

$$
\begin{aligned}
\mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x)dx = \int_a^b \frac{x}{b-a}dx = \frac{a+b}{2} \\
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_a^b \frac{x^2}{b-a}dx = \frac{a^2+ab+b^2}{3} \\
\mathrm{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{(b-a)^2}{12}.
\end{aligned}
$$

□

Example. (Quantization Error). Uniform distribution can be used to model quantization error in random signals. Let $X[n]$ be a discrete time signal at time n. ($X[n]$ has to be random, and independent at every time instant.) The quantized signal using a quantization step Δ is defined as

$$X_q[n] = \Delta \left[\frac{X[n]}{\Delta} \right],$$

where $[\cdot]$ denotes the rounding operator. The quantization error is the difference between the quantized sample and the true sample:

$$E_q[n] = X[n] - X_q[n].$$

The distribution of $E_q[n]$ is approximately uniform, with $E_q[n] \sim \text{Uniform}\left[-\frac{\Delta}{2}, \frac{\Delta}{2}\right]$. Then, one can show that the quantization error power is

$$P_e \stackrel{\text{def}}{=} \mathrm{Var}[E_q[n]] = \frac{\Delta^2}{12}. \tag{4.21}$$

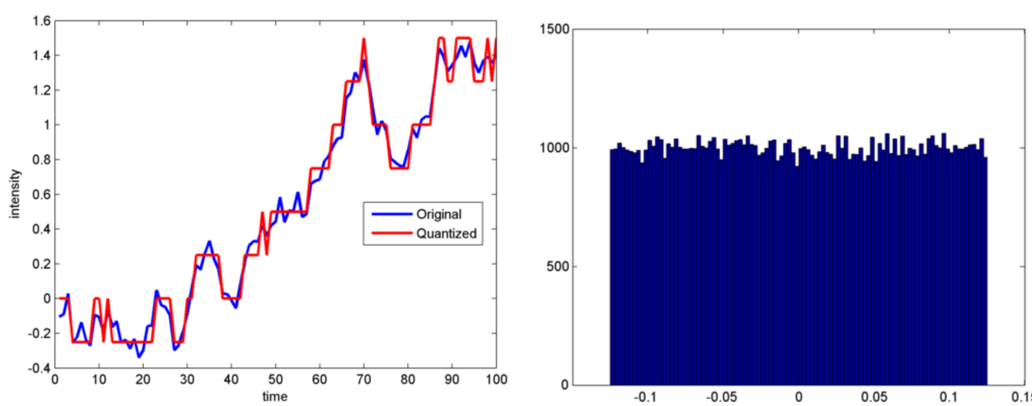

Figure 4.7: Illustration of quantization, and the histogram of the error.

Exponential Random Variable

Definition 11. *Let X be an exponential random variable. The PDF of X is*

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0, \\ 0, & \textit{otherwise}, \end{cases} \tag{4.22}$$

where $\lambda > 0$ is a parameter. We write

$$X \sim \text{Exponential}(\lambda)$$

to say that X is drawn from an exponential distribution of parameter λ.

The parameter λ of the exponential random variable determines the rate of decay. Thus, a large λ implies a faster decay.

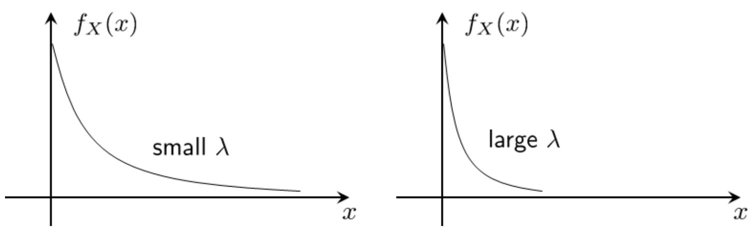

Figure 4.8: Exponential distribution

Theorem 7. *If $X \sim \text{Exponential}(\lambda)$, then*

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \textit{and} \quad \text{Var}[X] = \frac{1}{\lambda^2}.$$

Proof.

$$\begin{aligned} \mathbb{E}[X] &= \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^{\infty} \lambda x e^{-\lambda x} dx = -\int_0^{\infty} x d e^{-\lambda x} \\ &= -x e^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = \frac{1}{\lambda} \\ \mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^{\infty} \lambda x^2 e^{-\lambda x} dx = -\int_0^{\infty} x^2 d e^{-\lambda x} \\ &= -x^2 e^{-\lambda x}\Big|_0^{\infty} + \int_0^{\infty} 2x e^{-\lambda x} dx = 0 + \frac{2}{\lambda}\mathbb{E}[X] = \frac{2}{\lambda^2} \\ \text{Var}[X] &= \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{1}{\lambda^2}, \end{aligned}$$

where we used integration by parts in calculating $\mathbb{E}[X]$ and $\mathbb{E}[X^2]$. □

Example. (Relation to Poisson). An exponential random variable can be derived from a Poisson random variable, which is of important interest in studying traffics. Let N be the number of people arriving at a station. We assume that N is Poisson with rate λ. Let T be the inter-arrival time between two people. Then, for any duration t, the probability of observing n people is

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

Therefore,

$$\mathbb{P}[T > t] = \mathbb{P}[\text{inter-arrival time } > t] = \mathbb{P}[\text{no arrival in } t] = \mathbb{P}[N = 0] = e^{-\lambda t}.$$

Since $\mathbb{P}[T > t] = 1 - F_T(t)$, where $F_T(t)$ is the CDF of T, we can show that

$$f_T(t) = \frac{d}{dt} F_T(t) = \lambda e^{-\lambda t}. \tag{4.23}$$

That is, if N is Poisson with rate λ, then T is exponential with rate λ.

Gaussian Random Variable

Definition 12. *Let X be an Gaussian random variable. The PDF of X is*

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.24}$$

where (μ, σ^2) are parameters of the distribution. We write

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

to say that X is drawn from a Gaussian distribution of parameter (μ, σ^2).

Gaussian random variables are also called normal random variables. The parameters (μ, σ^2) are the mean and the variance, respectively. Note that a Gaussian random variable has a support from $-\infty$ to ∞. Thus, $f_X(x)$ has a non-zero value for any x, even though the value is extremely small. A Gaussian random variable is also symmetric about μ. If $\mu = 0$, then $f_X(x)$ is an even function.

Theorem 8. *If $X \sim \mathcal{N}(\mu, \sigma^2)$, then*

$$\mathbb{E}[X] = \mu, \quad \text{and} \quad \mathrm{Var}[X] = \sigma^2. \tag{4.25}$$

Proof. The expectation can be proved via substitution.

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \overset{(a)}{=} \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (y+\mu) e^{-\frac{y^2}{2\sigma^2}} dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} y e^{-\frac{y^2}{2\sigma^2}} dy + \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \mu e^{-\frac{y^2}{2\sigma^2}} dy \\
&\overset{(b)}{=} 0 + \mu \left(\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2\sigma^2}} dy \right) \overset{(c)}{=} \mu,
\end{aligned}$$

where in (a) we substitute $y = x - \mu$, in (b) we use the fact that the first integrand is odd so that the integration is 0, and in (c) we observe that integration over the entire sample space of the PDF yields 1.

The variance is also proved by substitution.

$$\begin{aligned}
\text{Var}[X] &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (x-\mu)^2 e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\
&\overset{(a)}{=} \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy, \quad \text{by letting} \quad y = \frac{x-\mu}{\sigma} \\
&= \frac{\sigma^2}{\sqrt{2\pi}} \left(-y e^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} \right) + \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = 0 + \sigma^2 \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \right) = \sigma^2,
\end{aligned}$$

where in (a) we substitute $y = (x-\mu)/\sigma$. □

Standard Gaussian

In many practical situations, we need to evaluate the probability $\mathbb{P}[a \leq X \leq b]$ of a Gaussian random variable X. This involves integration of the Gaussian PDF, i.e. determining the CDF. Unfortunately, there is no closed-form expression of $\mathbb{P}[a \leq X \leq b]$ in terms of (μ, σ^2). This leads to what we call the standard Gaussian.

Definition 13. *A* ***standard Gaussian*** *(or standard Normal) random variable X has a PDF*

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{4.26}$$

That is, $X \sim \mathcal{N}(0,1)$ is a Gaussian with $\mu = 0$ and $\sigma^2 = 1$.

Definition 14. *The* ***CDF*** *of the standard Gaussian is defined as the $\Phi(\cdot)$ function*

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{x^2}{2}} dx \tag{4.27}$$

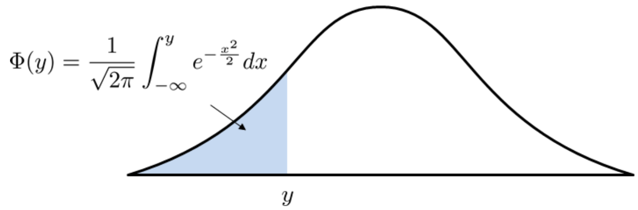

Figure 4.9: Definition of $\Phi(y)$.

Theorem 9 (CDF of an arbitrary Gaussian)**.** *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then,*

$$\mathbb{P}[X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right). \tag{4.28}$$

Proof. We start by expressing $\mathbb{P}[X \leq b]$:

$$\mathbb{P}[X \leq b] = \int_{-\infty}^{b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx.$$

Substituting $y = \frac{x-\mu}{\sigma}$, and using the definition of standard Gaussian, we have

$$\int_{-\infty}^{b} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \int_{-\infty}^{\frac{b-\mu}{\sigma}} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy = \Phi\left(\frac{b-\mu}{\sigma}\right).$$

□

Corollary 1. *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then,*

$$\mathbb{P}[a < X \leq b] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right). \tag{4.29}$$

Proof. Applying the Gaussian CDF twice yields the desired result:

$$\mathbb{P}[a < X \leq b] = \mathbb{P}[X \leq b] - \mathbb{P}[X \leq a] = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right).$$

□

Note that in the above corollary, the inequality signs of the two end points are not important. That is, the statement also holds for $\mathbb{P}[a \leq X \leq b]$ or $\mathbb{P}[a < X < b]$, because X is a continuous random variable at every x. Thus, $\mathbb{P}[X = a] = \mathbb{P}[X = b] = 0$ for any a and b.

Properties of $\Phi(y)$

Corollary 2. *Let $X \sim \mathcal{N}(\mu, \sigma^2)$. Then, the following results hold:*

- $\Phi(y) = 1 - \Phi(-y)$.
- $\mathbb{P}[X \geq b] = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$.
- $\mathbb{P}[|X| \geq b] = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right) + \Phi\left(\frac{-b-\mu}{\sigma}\right)$

Proof. Exercise. □

4.6 Function of Random Variables

One common question we encounter in practice is the transformation of random variables. The question is simple: Given a random variable X with PDF $f_X(x)$ and CDF $F_X(x)$, and suppose that $Y = g(X)$ for some function g, then what are $f_Y(y)$ and $F_Y(y)$?

General Principle

In general, this question can be answer by first looking at the CDF of Y:

$$F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[g(X) \leq y]. \tag{4.30}$$

If g is an invertible function, i.e., g^{-1} exists and gives an unique value, then

$$\mathbb{P}[g(X) \leq y] = \mathbb{P}[X \leq g^{-1}(y)] = F_X(g^{-1}(y)). \tag{4.31}$$

The CDF is the integration of the variable x

$$F_X(g^{-1}(y)) = \int_{-\infty}^{g^{-1}(y)} f_X(x')dx', \tag{4.32}$$

and hence by fundamental theorem of calculus, we have

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) \\
&= \frac{d}{dy} \int_{-\infty}^{g^{-1}(y)} f_X(x')dx' \\
&= \left(\frac{d\, g^{-1}(y)}{dy}\right) \cdot f_X(g^{-1}(y)).
\end{aligned} \tag{4.33}$$

When g^{-1} does not attain a unique value, e.g., $g(x) = x^2$ implies $g^{-1}(y) = \pm\sqrt{y}$. When this happens, then instead of writing $X \leq g^{-1}(y)$ we need to determine all possible X such that $g(X) \leq y$. The following examples will illustrate how we can do so.

Procedure.
To make the above discussion short we summarize them into two major steps:

- Step 1: Find the CDF $F_Y(y)$, which is $\mathbb{P}[Y \leq y]$, using $F_X(x)$.
- Step 2: Find the PDF $f_Y(y)$ by taking derivative on $F_Y(y)$.

Let us now consider a few examples.

Example 1. Let X be a random variable with PDF $f_X(x)$ and CDF $F_X(x)$. Let $Y = 2X+3$. Find $f_Y(y)$ and $F_Y(y)$. Express answers in terms of $f_X(x)$ and $F_X(x)$.

Solution. We first note that

$$F_Y(y) = \mathbb{P}[Y \leq y] = \mathbb{P}[2X+3 \leq y] = \mathbb{P}\left[X \leq \frac{y-3}{2}\right] = F_X\left(\frac{y-3}{2}\right).$$

Therefore, the PDF is

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}F_X\left(\frac{y-3}{2}\right) \\ &= F_X'\left(\frac{y-3}{2}\right)\frac{d}{dy}\left(\frac{y-3}{2}\right) = \frac{1}{2}f_X\left(\frac{y-3}{2}\right). \end{aligned}$$

Example 2. Let X be a random variable with PDF $f_X(x)$ and CDF $F_X(x)$. Suppose that $Y = X^2$, find $f_Y(y)$ and $F_Y(y)$. Express answers in terms of $f_X(x)$ and $F_X(x)$.

Solution. To solve this problem, we note that

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \leq y] = \mathbb{P}[X^2 \leq y] \\ &= \mathbb{P}[-\sqrt{y} \leq X \leq \sqrt{y}] \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Therefore, the PDF is

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right) \\ &= F_X'(\sqrt{y})\frac{d}{dy}\sqrt{y} - F_X'(-\sqrt{y})\frac{d}{dy}\left(-\sqrt{y}\right) \\ &= \frac{1}{2\sqrt{y}}\left(f_X(\sqrt{y}) + f_X(-\sqrt{y})\right). \end{aligned}$$

Example 3. Let $X \sim \text{Uniform}(0, 2\pi)$. Suppose $Y = \cos X$. Find $f_Y(y)$ and $F_Y(y)$.

Solution. First of all, we need to find the CDF of X. This can be done by noting that

$$F_X(x) = \int_{-\infty}^{x} f_X(x')dx' = \int_{0}^{x} \frac{1}{2\pi} dx' = \frac{x}{2\pi}.$$

Thus, the CDF of Y is

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \le y] = \mathbb{P}[\cos X \le y] \\ &= \mathbb{P}[\cos^{-1} y \le X \le 2\pi - \cos^{-1} y] \\ &= F_X(2\pi - \cos^{-1} y) - F_X(\cos^{-1} y) \\ &= 1 - \frac{\cos^{-1} y}{\pi}. \end{aligned}$$

The PDF of Y is

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(1 - \frac{\cos^{-1} y}{\pi}\right) \\ &= \frac{1}{\pi\sqrt{1-y^2}}, \end{aligned}$$

where we used the fact that $\frac{d}{dy}\cos^{-1} y = \frac{-1}{\sqrt{1-y^2}}$.

Example 4. Let X be a random variable with PDF

$$f_X(x) = ae^{x}e^{-ae^{x}}.$$

Let $Y = e^X$, and find $f_Y(y)$.

Solution. To solve this problem, we first note that

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \le y] = \mathbb{P}[e^X \le y] \\ &= \mathbb{P}[X \le \log y] \\ &= \int_{-\infty}^{\log y} ae^{x}e^{-ae^{x}} dx. \end{aligned}$$

To find the PDF, we recall the fundamental theorem of calculus. This will give us

$$\begin{aligned} f_Y(y) &= \frac{d}{dy}\int_{-\infty}^{\log y} ae^{x}e^{-ae^{x}} dx \\ &= \left(\frac{d}{dy}\log y\right)\left(\frac{d}{d\log y}\int_{-\infty}^{\log y} ae^{x}e^{-ae^{x}} dx\right) \\ &= \frac{1}{y} ae^{\log y} e^{-ae^{\log y}} \\ &= ae^{-ay}. \end{aligned}$$

4.7 Generating Random Numbers

For common types of random variables, e.g., Gaussian or exponential, there are often built-in functions to generate the random numbers. However, for arbitrary PDFs, how can we generate the random numbers?

Generating arbitrary random numbers can be done with the help of the following theorem.

Theorem 10. *Let X be a random variable with an invertible CDF $F_X(x)$, i.e., F_X^{-1} exists. If $Y = F_X(X)$, then $Y \sim \text{Uniform}(0,1)$.*

Proof. First, we know that if $U \sim \text{Uniform}(0,1)$, then $f_U(u) = 1$ for $0 \leq u \leq 1$ and so

$$F_U(u) = \int_{-\infty}^{u} f_U(u)du = u,$$

for $0 \leq u \leq 1$. If $Y = F_X(X)$, then the CDF of Y is

$$\begin{aligned} F_Y(y) &= \mathbb{P}[Y \leq y] = \mathbb{P}[F_X(X) \leq y] \\ &= \mathbb{P}[X \leq F_X^{-1}(y)] \\ &= F_X(F_X^{-1}(y)) = y. \end{aligned}$$

Therefore, we showed that the CDF of Y is the CDF of a uniform distribution. This completes the proof.

The consequence of the theorem is the following procedure in generating random numbers:

- Step 1: Generate a random number $U \sim \text{Uniform}(0,1)$.
- Step 2: Let $Y = F_X^{-1}(U)$. Then the distribution of Y is F_X.

Example. Suppose we want to draw random numbers from an exponential distribution. Then, we can follow the above two steps as

- Step 1: Generate a random number $U \sim \text{Uniform}(0,1)$.
- Step 2: Find F_X^{-1}. Since X is exponential, we know that $F_X(x) = 1 - e^{-\lambda x}$. This gives the inverse $F_X^{-1}(x) = -\frac{1}{\lambda}\log(1-x)$. Therefore, by letting $Y = F_X^{-1}(U)$, the distribution of Y will be exponential.