4.1.3 Functions of Continuous Random Variables

If X is a continuous random variable and $Y = g(X)$ is a function of X, then Y itself is a random variable. Thus, we should be able to find the CDF and PDF of Y. It is usually more straightforward to start from the CDF and then to find the PDF by taking the derivative of the CDF. Note that before differentiating the CDF, we should check that the CDF is continuous. As we will see later, the function of a continuous random variable might be a non-continuous random variable. Let's look at an example.

Example 4.7

Let X be a $Uniform(0, 1)$ random variable, and let $Y = e^X$.

    a. Find the CDF of Y.
    b. Find the PDF of Y.
    c. Find EY.

Solution

First, note that we already know the CDF and PDF of X. In particular,

$$F_X(x) = \begin{cases} 0 & \text{for } x < 0 \\ x & \text{for } 0 \leq x \leq 1 \\ 1 & \text{for } x > 1 \end{cases}$$

It is a good idea to think about the range of Y before finding the distribution. Since e^x is an increasing function of x and $R_X = [0, 1]$, we conclude that $R_Y = [1, e]$. So we immediately know that

$$F_Y(y) = P(Y \leq y) = 0, \qquad \text{for } y < 1,$$

$$F_Y(y) = P(Y \leq y) = 1, \qquad \text{for } y \geq e.$$

    a. To find $F_Y(y)$ for $y \in [1, e]$, we can write

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(e^X \leq y) \\ &= P(X \leq \ln y) \end{aligned}$$

$$= F_X(\ln y) = \ln y$$

To summarize

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 1 \\ \ln y & \text{for } 1 \le y < e \\ 1 & \text{for } y \ge e \end{cases}$$

b. The above CDF is a continuous function, so we can obtain the PDF of Y by taking its derivative. We have

$$f_Y(y) = F_Y'(y) = \begin{cases} \frac{1}{y} & \text{for } 1 \le y \le e \\ 0 & \text{otherwise} \end{cases}$$

Note that the CDF is not technically differentiable at points 1 and e, but as we mentioned earlier we do not worry about this since this is a continuous random variable and changing the PDF at a finite number of points does not change probabilities.

c. To find the EY, we can directly apply LOTUS,
$$E[Y] = E[e^X] = \int_{-\infty}^{\infty} e^x f_X(x)dx$$
$$= \int_0^1 e^x dx$$
$$= e - 1.$$

For this problem, we could also find EY using the PDF of Y,
$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y)dy$$
$$= \int_1^e y \frac{1}{y} dy$$
$$= e - 1.$$

Note that since we have already found the PDF of Y it did not matter which method we used to find $E[Y]$. However, if the problem only asked for $E[Y]$ without asking for the PDF of Y, then using LOTUS would be much easier.

Example 4.8
Let $X \sim Uniform(-1, 1)$ and $Y = X^2$. Find the CDF and PDF of Y.

Solution

First, we note that $R_Y = [0, 1]$. As usual, we start with the CDF. For $y \in [0, 1]$, we have

$$\begin{aligned} F_Y(y) &= P(Y \le y) \\ &= P(X^2 \le y) \\ &= P(-\sqrt{y} \le X \le \sqrt{y}) \\ &= \frac{\sqrt{y} - (-\sqrt{y})}{1 - (-1)} \qquad\qquad \text{since } X \sim Uniform(-1, 1) \\ &= \sqrt{y}. \end{aligned}$$

Thus, the CDF of Y is given by

$$F_Y(y) = \begin{cases} 0 & \text{for } y < 0 \\ \sqrt{y} & \text{for } 0 \le y \le 1 \\ 1 & \text{for } y > 1 \end{cases}$$

Note that the CDF is a continuous function of Y, so Y is a continuous random variable. Thus, we can find the PDF of Y by differentiating $F_Y(y)$,

$$f_Y(y) = F_Y'(y) = \begin{cases} \frac{1}{2\sqrt{y}} & \text{for } 0 \le y \le 1 \\ 0 & \text{otherwise} \end{cases}$$

The Method of Transformations

So far, we have discussed how we can find the distribution of a function of a continuous random variable starting from finding the CDF. If we are interested in finding the PDF of $Y = g(X)$, and the function g satisfies some properties, it might be easier to use a method called the method of transformations. Let's start with the case where g is a function satisfying the following properties:

- $g(x)$ is differentiable;
- $g(x)$ is a strictly increasing function, that is, if $x_1 < x_2$, then $g(x_1) < g(x_2)$.

Now, let X be a continuous random variable and $Y = g(X)$. We will show that you can directly find the PDF of Y using the following formula.

$$f_Y(y) = \begin{cases} \frac{f_X(x_1)}{g'(x_1)} = f_X(x_1) \cdot \frac{dx_1}{dy} & \text{where } g(x_1) = y \\ 0 & \text{if } g(x) = y \text{ does not have a solution} \end{cases}$$

Note that since g is strictly increasing, its inverse function g^{-1} is well defined. That is, for each $y \in R_Y$, there exists a unique x_1 such that $g(x_1) = y$. We can write $x_1 = g^{-1}(y)$.

$$F_Y(y) = P(Y \le y)$$

$$= P(g(X) \leq y)$$
$$= F_X(g^{-1}(y)).$$

To find the PDF of Y, we differentiate

$$f_Y(y) = \frac{d}{dy}F_X(x_1) \qquad\qquad \text{where } g(x_1) = y$$
$$= \frac{dx_1}{dy} \cdot F_X'(x_1)$$
$$= f_X(x_1)\frac{dx_1}{dy}$$
$$= \frac{f_X(x_1)}{g'(x_1)} \qquad\qquad \text{since } \frac{dx}{dy} = \frac{1}{\frac{dy}{dx}}.$$

We can repeat the same argument for the case where g is strictly decreasing. In that case, $g'(x_1)$ will be negative, so we need to use $|g'(x_1)|$. Thus, we can state the following theorem for a strictly monotonic function. (A function $g : \mathbb{R} \to \mathbb{R}$ is called strictly monotonic if it is strictly increasing or strictly decreasing.)

Theorem 4.1

Suppose that X is a continuous random variable and $g : \mathbb{R} \to \mathbb{R}$ is a strictly monotonic differentiable function. Let $Y = g(X)$. Then the PDF of Y is given by

$$f_Y(y) = \begin{cases} \frac{f_X(x_1)}{|g'(x_1)|} = f_X(x_1).|\frac{dx_1}{dy}| & \text{where } g(x_1) = y \\ 0 & \text{if } g(x) = y \text{ does not have a solution} \end{cases} \qquad (4.5)$$

To see how to use the formula, let's look at an example.

Example 4.9

Let X be a continuous random variable with PDF

$$f_X(x) = \begin{cases} 4x^3 & 0 < x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

and let $Y = \frac{1}{X}$. Find $f_Y(y)$.

Solution

First note that $R_Y = [1, \infty)$. Also, note that $g(x)$ is a strictly decreasing and differentiable function on $(0, 1]$, so we may use Equation 4.5. We have $g'(x) = -\frac{1}{x^2}$.

For any $y \in [1, \infty)$, $x_1 = g^{-1}(y) = \frac{1}{y}$. So, for $y \in [1, \infty)$

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|}$$
$$= \frac{4x_1^3}{|-\frac{1}{x_1^2}|}$$
$$= 4x_1^5$$
$$= \frac{4}{y^5}.$$

Thus, we conclude

$$f_Y(y) = \begin{cases} \frac{4}{y^5} & y \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

Theorem 4.1 can be extended to a more general case. In particular, if g is not monotonic, we can usually divide it into a finite number of monotonic differentiable functions. Figure 4.3 shows a function g that has been divided into monotonic parts. We may state a more general form of Theorem 4.1.
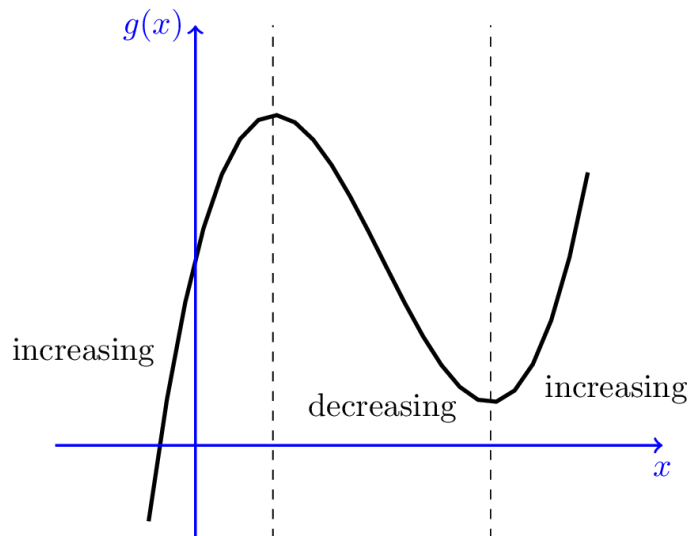

Fig.4.3 - Partitioning a function to monotone parts.

Theorem 4.2

Consider a continuous random variable X with domain R_X, and let $Y = g(X)$. Suppose that we can partition R_X into a finite number of intervals such that $g(x)$ is strictly monotone and differentiable on each partition. Then the PDF of Y is given by

$$f_Y(y) = \sum_{i=1}^{n} \frac{f_X(x_i)}{|g'(x_i)|} = \sum_{i=1}^{n} f_X(x_i) \cdot \left| \frac{dx_i}{dy} \right| \qquad (4.6)$$

where $x_1, x_2, \ldots, x_n$ are real solutions to $g(x) = y$.

Let us look at an example to see how we can use Theorem 4.2.

Example 4.10

Let X be a continuous random variable with PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \text{for all } x \in \mathbb{R}$$

and let $Y = X^2$. Find $f_Y(y)$.

Solution

We note that the function $g(x) = x^2$ is strictly decreasing on the interval $(-\infty, 0)$, strictly increasing on the interval $(0, \infty)$, and differentiable on both intervals, $g'(x) = 2x$. Thus, we can use Equation 4.6. First, note that $R_Y = (0, \infty)$. Next, for any $y \in (0, \infty)$ we have two solutions for $y = g(x)$, in particular,

$$x_1 = \sqrt{y}, \quad x_2 = -\sqrt{y}$$

Note that although $0 \in R_X$ it has not been included in our partition of R_X. This is not a problem, since $P(X = 0) = 0$. Indeed, in the statement of Theorem 4.2, we could replace R_X by $R_X - A$, where A is any set for which $P(X \in A) = 0$. In particular, this is convenient when we exclude the endpoints of the intervals. Thus, we have

$$f_Y(y) = \frac{f_X(x_1)}{|g'(x_1)|} + \frac{f_X(x_2)}{|g'(x_2)|}$$

$$= \frac{f_X(\sqrt{y})}{|2\sqrt{y}|} + \frac{f_X(-\sqrt{y})}{|-2\sqrt{y}|}$$

$$= \frac{1}{2\sqrt{2\pi y}} e^{-\frac{y}{2}} + \frac{1}{2\sqrt{2\pi y}} e^{-\frac{y}{2}}$$

$$= \frac{1}{\sqrt{2\pi y}} e^{-\frac{y}{2}}, \text{ for } y \in (0, \infty).$$