Module 13
Transformations of Absolutely Continuous Random Variables

- X: an A.C. r.v. with d.f. $F_X(\cdot)$ and p.d.f. $f_X(\cdot)$;
- $$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \ x \in \mathbb{R};$$
- For $-\infty < a < b < \infty$, $P(\{X = a\}) = 0$,

$$P(\{a < X \leq b\}) = P(\{a < X < b\}) = P(\{a \leq X < b\})$$

$$= P(\{a \leq X \leq b\}) = \int_{a}^{b} f_X(t)dt;$$

- In general, for any set $A \subseteq \mathbb{R}$,

$$P(\{X \in A\}) = \int_{A} f_X(t)dt = \int_{-\infty}^{\infty} f_X(t)I_A(t)dt.$$

- We will assume throughout that $f_X(\cdot)$ is continuous everywhere except at (possibly) finite number of points (say, $x_1, x_2, \ldots, x_n$), where it has jump discontinuities. In that case $F_X(\cdot)$ is differentiable everywhere except at discontinuity points $\{x_1, \ldots, x_n\}$ of $f_X(\cdot)$. Moreover

$$f_X(t) = \begin{cases} F'_X(t), & \text{if } t \notin \{x_1, \ldots, x_n\} \\ \\ 0, & \text{otherwise} \end{cases}.$$

- Support

$$S_X = \{x \in \mathbb{R} : F_X(x+\epsilon) - F_X(x-\epsilon) > 0, \ \forall \ \epsilon > 0\}.$$

- $g : \mathbb{R} \to \mathbb{R}$: a given function;
- Then $Y = g(X)$ is a r.v.;
- **Goal**: To find the probability distribution (i.e., d.f. $F_Y(\cdot)$ and/or, p.d.f./p.m.f. $f_Y(\cdot)$) of $Y = g(X)$;

Remark 1: We have seen that when X is discrete, $Y = g(X)$ is also discrete. When X is A.C., $Y = g(X)$ may not be A.C. (or even continuous) as the following example illustrates.

Example 1: Let X be an A.C. r.v. with p.d.f.

$$f_X(x) = \begin{cases} \frac{1}{2}, & \text{if } -1 < x < 1 \\ 0, & \text{otherwise} \end{cases}.$$

Let $Y = [X]$ (maximum integer contained in X). Note that $P(\{X \in (-1, 1)\}) = 1$.

$$Y = \begin{cases} -1, & \text{if } -1 < X < 0 \\ 0, & \text{if } 0 \leq X < 1 \end{cases}.$$

Then

$$\begin{aligned} P(\{Y = -1\}) &= P(\{-1 < X < 0\}) = \int_{-1}^{0} f_X(x)dx \\ &= \int_{-1}^{0} \frac{1}{2}dx = \frac{1}{2}, \end{aligned}$$

$$
\begin{aligned}
P(\{Y=0\}) &= P(\{0 \le X < 1\}) = \int_0^1 f_X(x)dx \\
&= \int_0^1 \frac{1}{2}dx = \frac{1}{2}.
\end{aligned}
$$

Thus Y is discrete with support $S_Y = \{-1, 0\}$ and p.m.f.

$$
f_Y(y) = \begin{cases} \frac{1}{2}, & \text{if } y \in \{-1, 0\} \\ 0, & \text{otherwise} \end{cases}.
$$

The following result provides sufficient conditions under which a function of an A.C. random variable is A.C.

Result 1: Suppose $S_X = \bigcup_{i=1}^{k} S_{i,X}$, where $\{S_{i,X}, i = 1, \ldots, k\}$ is a collection of disjoint intervals and in $S_{i,X}$ $(i = 1, ..., k)$, $g : S_{i,X} \to \mathbb{R}$ is strictly monotone with inverse function $g_i^{-1}(y)$ such that $\frac{d}{dy} g_i^{-1}(y)$ is continuous. Let $g(S_{i,X}) = \{g(x) : x \in S_{i,X}\}$, $i = 1, ..., k$. Then the r.v. $Y = g(X)$ is A.C. with p.d.f.

$$f_Y(y) = \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right) \left| \frac{d}{dy} g_i^{-1}(y) \right| I_{g(S_{i,X})}(y),$$

where, for a set A, $I_A(\cdot)$ denotes its indicator function, i.e.,

$$I_A(y) = \begin{cases} 1, & \text{if } y \in A \\ 0, & \text{otherwise} \end{cases}.$$

Corollary 1: Suppose that $g : S_X \to \mathbb{R}$ is strictly monotone with inverse function $g^{-1}(y)$ such that $\frac{d}{dy}g^{-1}(y)$ is continuous. Let $g(S_X) = \{g(x) : x \in S_X\}$. Then $Y = g(X)$ is of A.C. type with p.d.f.

$$f_Y(y) = f_X(g^{-1}(y)) \left| \frac{d}{dy} g^{-1}(y) \right| I_{g(S_X)}(y).$$

Example 2: Let X be an A.C. r.v. with p.d.f.

$$f_X(x) = \begin{cases} \frac{|x|}{2}, & \text{if } -1 < x < 1 \\ \frac{x}{3}, & \text{if } 1 < x < 2 \\ 0, & \text{otherwise} \end{cases},$$

and let $Y = X^2$.

(a) Find the p.d.f. of Y and hence find the d.f. of Y;

(b) Find the d.f. of Y and hence find the p.d.f. of Y.

Solution: $S_X = [-1, 2] = [-1, 0) \cup [0, 2] = S_{1,X} \cup S_{2,X}$.
$g(x) = x^2,\ x \in S_X$, is monotone in $S_{1,X}$ and $S_{2,X}$.

$S_{1,X} = [-1, 0)$	$S_{2,X} = [0, 2]$
$g_1^{-1}(y) = -\sqrt{y}$	$g_2^{-1}(y) = \sqrt{y}$
$\frac{d}{dy} g_1^{-1}(y) = \frac{-1}{2\sqrt{y}}$	$\frac{d}{dy} g_2^{-1}(y) = \frac{1}{2\sqrt{y}}$
$g(S_{1,X}) = (0, 1]$	$g(S_{2,X}) = [0, 4)$
$y \in g(S_{1,X}) \Leftrightarrow 0 < y \leq 1$	$y \in g(S_{2,X}) \Leftrightarrow 0 \leq y \leq 4$
$f_X(g_1^{-1}(y)) \left\lvert \frac{d}{dy} g_1^{-1}(y) \right\rvert$	$f_X(g_2^{-1}(y)) \left\lvert \frac{d}{dy} g_2^{-1}(y) \right\rvert$
$= f_X(-\sqrt{y}) \left\lvert \frac{-1}{2\sqrt{y}} \right\rvert I_{(0,1]}(y)$	$= f_X(\sqrt{y}) \left\lvert \frac{1}{2\sqrt{y}} \right\rvert I_{[0,4]}(y)$

Thus a p.d.f. of Y is

$$\begin{aligned} f_Y(y) &= \sum_{i=1}^{2} f_X(g_i^{-1}(y)) \left\lvert \frac{d}{dy} g_i^{-1}(y) \right\rvert I_{g(S_{i,X})}(y) \\ &= \begin{cases} \frac{1}{2}, & \text{if } 0 < y < 1 \\ \frac{1}{6}, & \text{if } 1 \leq y < 4 \\ 0, & \text{otherwise} \end{cases} . \end{aligned}$$

$$F_Y(y) = P(\{Y \le y\}) = \int_{-\infty}^{y} f_Y(t)dt = \begin{cases} 0, & \text{if } y < 0 \\ \\ 1, & \text{if } y \ge 4 \end{cases}.$$

For $0 \le y < 1$,

$$F_Y(y) = P(\{Y \le y\}) = \int_0^y \frac{1}{2} dt = \frac{y}{2}.$$

For $1 \le y < 4$

$$F_Y(y) = \int_0^1 \frac{1}{2} dt + \int_1^y \frac{1}{6} dt = \frac{y+2}{6}.$$

Thus the d.f. of Y is

$$F_Y(y) = \begin{cases} 0, & \text{if } y < 0 \\ \frac{y}{2}, & \text{if } 0 \le y < 1 \\ \frac{y+2}{6}, & \text{if } 1 \le y < 4 \\ 1, & \text{if } y \ge 4 \end{cases}.$$

(b) For $y < 0$,

$$F_Y(y) = P(\{Y \leq y\}) = P(\{X^2 \leq y\}) = 0.$$

For $y \geq 0$,

$$F_Y(y) = P(\{X^2 \leq y\}) = P(\{-\sqrt{y} \leq X \leq \sqrt{y}\}) = \int_{-\sqrt{y}}^{\sqrt{y}} f_X(t)dt.$$

For $0 \leq y < 1$,

$$F_Y(y) = \int_{-\sqrt{y}}^{\sqrt{y}} \frac{|t|}{2} dt = \frac{y}{2}.$$

For $1 \leq y < 4$,

$$\begin{aligned} F_Y(y) &= \int_{-1}^{1} \frac{|t|}{2} dt + \int_{1}^{\sqrt{y}} \frac{t}{3} dt \\ &= \frac{y+2}{6}. \end{aligned}$$

For $y \geq 4$, $F_Y(y) = 1$.

Thus the d.f. of Y is

$$F_Y(y) = \begin{cases} 0, & \text{if } y < 0 \\ \frac{y}{2}, & \text{if } 0 \leq y < 1 \\ \frac{y+2}{6}, & \text{if } 1 \leq y < 4 \\ 1, & \text{if } y \geq 4 \end{cases}.$$

Clearly $S_Y = [0, 4]$, $F_Y(\cdot)$ is differentiable everywhere except at points $0, 1$ and 4. Let

$$g(y) = \begin{cases} F'_Y(y), & \text{if } y \notin \{0, 1, 4\} \\ 0, & \text{otherwise} \end{cases}$$

$$= \begin{cases} \frac{1}{2}, & \text{if } 0 < y < 1 \\ \\ \frac{1}{6}, & \text{if } 1 < y < 4 \\ \\ 0, & \text{otherwise} \end{cases}.$$

Then $\int_{-\infty}^{\infty} g(y)dy = 1$. Thus Y is of A.C. type with p.d.f.

$$f_Y(y) = g(y) = \begin{cases} \frac{1}{2}, & \text{if } 0 < y < 1 \\ \\ \frac{1}{6}, & \text{if } 1 < y < 4 \\ \\ 0, & \text{otherwise} \end{cases}.$$

Take home problem

Let X be a r.v. with p.d.f.

$$f_X(x) = \begin{cases} e^{-x}, & \text{if } x > 0 \\ 0, & \text{otherwise} \end{cases}.$$

Let $Y = 2X + 3$.

(a) Find the p.d.f. of Y and hence find the d.f. of Y;

(b) Find the d.f. of Y and hence find the p.d.f. of Y.

Abstract of Next Module

- X: a r.v. associated with a random experiment $\mathcal{E}$;
- Each time the random experiment is performed we get a value of X;

Question: If the random experiment is performed infinitely what is the mean (or expectation) of observed values of X or $g(X)$, for some function real-valued function $g(\cdot)$?
In the discrete case, the relative frequency interpretation of probability suggests that we take

$$\begin{aligned} E(g(X)) &= \lim_{N\to\infty} \frac{\sum_{x\in S_X} g(x) \times \text{ Number of times we get } \{X=x\}}{N} \\ &= \sum_{x\in S_X} g(x) \lim_{N\to\infty} \frac{\text{frequency of } \{X=x\}}{N} \\ &= \sum_{x\in S_X} g(x) P(\{X=x\}) = \sum_{x\in S_X} g(x) f_X(x). \end{aligned}$$