2. Random variables and their distributions

Frequently, when an experiment is performed, we are interested mainly in some function of the outcome as opposed to the actual outcome itself. For instance, in tossing dice, we are often interested in the sum of the two dice and are not really concerned about the separate values of each die.

Definition 2.1 (**Random variable).** Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. A real-valued function $X : \Omega \to \mathbb{R}$ is said to be random variable if for any $B \in \mathcal{B}(\mathbb{R})$, $X^{-1}(B) \in \mathcal{F}$.

Example 2.1. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $A \in \mathcal{F}$. Define a function $X : \Omega \to \mathbb{R}$ by*

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A \end{cases}$$

Then X is a random variable. Indeed, for any $x \in \mathbb{R}$, we see that

$$\{\omega \in \Omega : X(\omega) \le x\} = \begin{cases} \emptyset & \text{if } x < 0 \\ A^{\complement} & \text{if } 0 \le x < 1 \\ \Omega & \text{if } x \ge 1. \end{cases}$$

The function X is called an indicator function of A and often denoted by χ_A or $\mathbf{I}_A$.

Example 2.2. *Let $\Omega = \{0, 1, 2\}$ and $\mathcal{F} = \{\emptyset, \{0\}, \{1, 2\}, \Omega\}$. Then $(\Omega, \mathcal{F})$ is a measurable space. Define a function X on Ω by $X(i) = i$ for $i \in \Omega$. Then X is NOT a random variable. Indeed, for any $x \in \mathbb{R}$, we see that*

$$\{i \in \Omega : X(i) \le x\} = \begin{cases} \emptyset & \text{if } x < 0 \\ \{0\} & \text{if } 0 \le x < 1 \\ \{0,1\} & \text{if } 1 \le x < 2 \\ \Omega & \text{if } x \ge 2. \end{cases}$$

Since $\{0, 1\} \notin \mathcal{F}$, by definition X is NOT a random variable.

For any $B \in \mathcal{B}(\mathbb{R})$, we are interested in $\mathbb{P}(X \in B)$. Let X be a random variable defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Define a function P_X on $\mathcal{B}(\mathbb{R})$ via

$$P_X(B) := \mathbb{P}(X \in B)$$

One can easily check that P_X is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. This is called distribution of X.

Definition 2.2 (**Distribution function/ Cumulative distribution function (cdf)).** Let X be a random variable defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the function $F_X : \mathbb{R} \to [0, 1]$ defined by

$$F_X(x) = \mathbb{P}(X \le x) = P_X((-\infty, x])$$

is called the distribution function or cumulative distribution function (cdf) of the random variable X.

Example 2.3. *A fair coin is tossed twice: $\Omega = \{HH, HT, TH, TT\}$. Take $\mathcal{F} = \mathcal{P}(\Omega)$ and define a probability measure $\mathcal{P}(A) = \frac{|A|}{4}$ for any $A \in \mathcal{F}$. For $\omega \in \Omega$, let $X(\omega)$ be the number of heads so that*

$$X(HH) = 2, \quad X(HT) = X(TH) = 1, \quad X(TT) = 0.$$

Then X is a random variable. Indeed for any $x \in \mathbb{R}$, one has

$$\{\omega : X(\omega) \leq x\} = \begin{cases} \emptyset & \text{for } x < 0 \\ \{TT\} & \text{for } 0 \leq x < 1 \\ \{TT, HT, TH\} & \text{for } 1 \leq x < 2 \\ \Omega & \text{for } x \geq 2. \end{cases}$$

The distribution function F_X of X is then given by

$$F_X(x) = \mathbb{P}(\omega : X(\omega) \leq x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{1}{4} & \text{for } 0 \leq x < 1 \\ \frac{3}{4} & \text{for } 1 \leq x < 2 \\ 1 & \text{for } x \geq 2. \end{cases}$$

Next we discuss some essential properties of distribution function.

Lemma 2.1. *The distribution function F_X satisfies the following properties:*

a) **Nondecreasing:** *if $x < y$, then $F_X(x) \leq F_X(y)$.*
b) **Right continuity:** *F is right-continuous i.e., $F_X(x+h) \to F_X(x)$ as $h \downarrow 0$.*
c) **Left limit:** *$F_X(\cdot)$ has left limit and $F_X(x-) = \mathbb{P}(X < x)$.*
d) *$\lim_{x\to\infty} F_X(x) = 1$ and $\lim_{x\to-\infty} F_X(x) = 0$.*
e) *$\mathbb{P}(x < X \leq y) = F_X(y) - F_X(x)$.*
f) *$\mathbb{P}(x \leq X \leq y) = F_X(y) - F_X(x-)$.*
g) *$\mathbb{P}(x \leq X < y) = F_X(y-) - F_X(x-)$, $\mathbb{P}(x < X < y) = F_X(y-) - F_X(x)$.*
h) *$\mathbb{P}(X = x) = F_X(x) - F_X(x-)$.*

Proof. **Proof of** a)**:** For $x < y$, we have $\{X \leq x\} \subseteq \{X \leq y\}$ and hence $F_X(x) \leq F_X(y)$.

Proof of b)**:** Since F is nondecreasing, it is sufficient to show that $F_X(x_n) \to F_X(x)$ for any sequence of numbers $x_n \downarrow x$ with $x_1 \geq x_2 \geq \ldots x_n > x$. Define $A_n := \{X \leq x_n\}$. Then $A_{n+1} \subseteq A_n$ and $\cap_{n=1}^{\infty} A_n = \{X \leq x\}$. Hence by using the property of $\mathbb{P}$, we get

$$F_X(x) = \mathbb{P}(X \leq x) = \mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} F_X(x_n).$$

Proof of c)**:** Let $x \in \mathbb{R}$ be fixed. Let $\{x_n\}$ be such that $x_1 \leq x_2 \leq \ldots < x$ and $\lim_{n\to\infty x_n = x}$. Take $A_n = \{X \leq x_n\}$. Then $A_n \subseteq A_{n+1}$ and $\cap_{n=1}^{\infty} A_n = \{X < x\}$. Thus, we have

$$\mathbb{P}(X < x) = \mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} F_X(x_n).$$

Thus, $F_X(\cdot)$ has left limit and $F_X(x-) = \mathbb{P}(X < x)$.

Proof of d)**:** Observe that $\{X \leq n\} \subseteq \{X \leq (n+1)\}$ and $\Omega = \cup_{n=1}^{\infty}\{X \leq n\}$. Hence by using the property of $\mathbb{P}$, we get

$$1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{n=1}^{\infty}\{X \leq n\}) = \lim_{n\to\infty} \mathbb{P}(\{X \leq n\}) = \lim_{n\to\infty} F_X(n) = \lim_{x\to\infty} F_X(x).$$

For the second part take $A_n = \{X \leq -n\}$. Then A_n is decreasing and $\cap_{n=1}^{\infty} A_n = \emptyset$. Hence one has

$$0 = \mathbb{P}(\emptyset) = \mathbb{P}(\cap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mathbb{P}(A_n) = \lim_{n\to\infty} F_X(-n) = \lim_{x\to-\infty} F_X(x).$$

Proof of e)**:** Since $\mathbb{P}_X$ is a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, by using the property of a measure that for any $B \subseteq A$, there holds $\mathbb{P}_X(A) - \mathbb{P}_X(B) = \mathbb{P}_X(A \setminus B)$ and the definition $F_X(x) = \mathbb{P}_X((-\infty, x])$, we see that

$$F_X(y) - F_X(x) = \mathbb{P}_X((x, y]) = \mathbb{P}(x < X \leq y).$$

Proof of f)**:** By e), $\mathbb{P}_X((x-\frac{1}{n}, y]) = F_X(y) - F_X(x-\frac{1}{n})$. Observe that the sequence of intervals $(x-\frac{1}{n}, y]$ decreases to $[x,y]$. Hence

$$\mathbb{P}(x \le X \le y) = \mathbb{P}_X([x,y]) = \lim_{n\to\infty} \mathbb{P}_X((x-\frac{1}{n}, y]) = \lim_{n\to\infty} \big(F_X(y) - F_X(x-\frac{1}{n})\big) = F_X(y) - F_X(x-).$$

Proof of g)**:** Note that $[x,y) = \cap_n A_n$ where $A_n = [x, y-\frac{1}{n}]$ and A_n are nondecreasing. Thus,

$$\mathbb{P}_X([x,y)) = \lim_{n\to\infty} \mathbb{P}_X A_n = F_X(y-) - F_X(x-).$$

For the second part, we take $A_n = (x - y - \frac{1}{n}]$. Then $A_n \uparrow (x,y)$ and hence we see that

$$\mathbb{P}(x < X < y) = \mathbb{P}_X((x,y)) = \lim_{n\to\infty} \mathbb{P}_X(A_n) = \lim_{n\to\infty} \big(F_X(y-\frac{1}{n}) - F_X(x)\big) = F_X(y-) - F_X(x).$$

Proof of h)**:** We see that, by using c) and the property of $\mathbb{P}_X$

$$\begin{aligned}\mathbb{P}(X = x) &= \mathbb{P}_X\big((-\infty, x] \setminus (-\infty, x)\big) = \mathbb{P}_X((-\infty,x]) - \mathbb{P}_X((-\infty,x)) \\ &= F_X(x) - \mathbb{P}(X < x) = F_X(x) - F_X(x-).\end{aligned}$$

□

2.1. **Discrete and Continuous random variables:**

Definition 2.3 (**Discrete random variable).** Let X be a real-valued random variable. It is said to be a discrete random variable if it takes the values in some countable subset $E = \{x_1, x_2, \ldots\}$, i.e., $\mathbb{P}(X \in E) = 1$.

Definition 2.4 (**Probability mass function (pmf)).** Let X be a discrete random variable, defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in a countable set E of $\mathbb{R}$. The function $p_X : \mathbb{R} \to [0,1]$ defined by

$$p_X(x) = \begin{cases} \mathbb{P}(X = x) & \text{if } x = x_i \in E \\ 0 & \text{otherwise} \end{cases}$$

is called probability mass function of X.

Let $p_i = \mathbb{P}(X = x_i)$. Then $p_i \ge 0$ and $\sum_i p_i = \sum_{i=1}^\infty \mathbb{P}(X = x_i) = \mathbb{P}\big(\cup_{i=1}^\infty \{X = x_i\}\big) = \mathbb{P}(X \in E) = 1$. The cdf of a discrete random variable may be given in terms of pmf:

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}(\cup_{x_i \le x}\{X = x_i\}) = \sum_{x_i \le x} \mathbb{P}(X = x_i) = \sum_{x_i \le x} p_X(x_i).$$

Example 2.4. *A fair coin is tossed two times. Let X be the number of heads obtained. Then clearly X is a discrete random variable. Its pmf is given by*

$$\begin{aligned} &p_X(0) = \mathbb{P}(X = 0) = \frac{1}{4}, \quad p_X(1) = \mathbb{P}(X = 2) = \frac{2}{4} = \frac{1}{2} \\ &p_X(2) = \mathbb{P}(X = 2) = \frac{1}{4}, \quad p_X(x) = 0, \quad \forall x \in \mathbb{R} \setminus \{0,1,2\}. \end{aligned}$$

Let us determine the following:

$$\begin{aligned} &\mathbb{P}(0.5 < X \le 4) = \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = \frac{3}{4}, \quad \mathbb{P}(-1.5 \le X < 1) = \mathbb{P}(X = 0) = \frac{1}{4}, \\ &\mathbb{P}(X \le 2) = \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) = 1. \end{aligned}$$

Definition 2.5 (**Continuous random variable).** Let X be a random variable with cdf F_X. Then X is said to be continuous if there exists a non-negative integrable function f_X such that

$$F_X(x) = \int_{-\infty}^x f_X(t)\, dt.$$

Observe that, since $\lim_{x\to\infty} F_X(x) = 1$, we have $\int_{-\infty}^{\infty} f_X(t)\,dt = 1$. Moreover, for any $a, b \in \mathbb{R}$ with $a < b$, we have

$$\mathbb{P}(a < X \leq b) = F_X(b) - F_X(a) = \int_{-\infty}^{b} f_X(t)\,dt - \int_{-\infty}^{a} f_X(t)\,dt = \int_{a}^{b} f_X(t)\,dt.$$

Notice that for any continuous random variable X, the cdf F_X is continuous and hence

i) $\mathbb{P}(X = a) = F_X(a) - F_X(a-) = 0.$

ii) $\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) = \int_a^b f_X(t)\,dt.$

Furthermore, for any $B \in \mathcal{B}(\mathbb{R})$, we obtain $\mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(t)\,dt.$

Definition 2.6 (Probability density function (pdf)). The function f_X given in Definition 2.5 is called probability density function of the random variable X.

Observe that if f_X is continuous, then F_X is differentaible at every point of x and $f_X = \frac{d}{dx}F_X$. Since F_X is absolutely continuous, it is differentiable at all x except a countably many points. For the points where F_X is not differentaible, we define $\frac{d}{dx}F_X = 0$. In this way, we conclude that

$$f_X(x) = \frac{d}{dx}F_X(x) \quad \forall\ x.$$

Example 2.5. *Let X be a random variable whose cdf is given by*

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

We like to determine the pdf of X and then calculate $\mathbb{P}(X \leq \frac{1}{3})$. Observe that F_X is continuous on $\mathbb{R}$ and differentiable except the zero point. Thus

$$f_X(x) = \begin{cases} xe^{-x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

By the definition of cdf, we get that

$$\mathbb{P}(x \leq \frac{1}{3}) = F_X(\frac{1}{3}) = 1 - \frac{4}{3}e^{-\frac{1}{3}}.$$

Example 2.6. *Let X be a continuous random variable with pdf*

$$f_X(x) = \begin{cases} kx(1-x), & 0 < x < 1 \\ 0, & otherwise \end{cases}$$

We would like to determine k and then calculate $\mathbb{P}(X > 0.3)$. Since f_X is pdf, it must satisfy the condition $\int_{\mathbb{R}} f_X(x)\,dx = 1$. This implies that $k = 6$. Using the definition of continuous random variable, we get

$$\mathbb{P}(X > 0.3) = 1 - \mathbb{P}(X \leq 0.3) = 1 - F_X(0.3) = 1 - 6\int_0^{0.3} x(1-x)\,dx = 0.784.$$

2.2. **Functions of a random variables:** Let X be a random variable defined on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let g be a Borel-measurable function on $\mathbb{R}$. Then $Y = g(X)$ is a random variable. Indeed, for any $y \in \mathbb{R}$,

$$\{\omega : Y(\omega) \leq y\} = \{\omega : X \in g^{-1}((-\infty, y])\}$$

Since g is Borel-measurable, $g^{-1}((-\infty, y]) \in \mathcal{B}(\mathbb{R})$ and hence $\{\omega : X \in g^{-1}((-\infty, y])\} \in \mathcal{F}$.

We are interested in determining the distribution function of $Y = g(X)$ in terms of the cdf F_X of X. It is given by

$$F_Y(y) = \mathbb{P}(g(X) \leq y) = \mathbb{P}(X \in g^{-1}((-\infty, y])).$$

Example 2.7. *Let X be a random variable with cdf F_X. Take $Y = |X|$. Then*

$$F_Y(y) = \mathbb{P}(-y \le X \le y) = F_X(y) - F_X(-y) + \mathbb{P}(X = -y).$$

If X is a continuous random variable, then $\mathbb{P}(X = -y) = 0$ and hence

$$F_Y(y) = \begin{cases} F_X(y) - F_X(-y), & y > 0 \\ 0, & y \ge 0. \end{cases}$$

If X is discrete random variable, then

$$F_Y(y) = \begin{cases} F_X(y) - F_X(-y) + \mathbb{P}(X = -y), & y > 0 \\ 0, & y \ge 0. \end{cases}$$

Example 2.8. *Suppose that X is a continuous random variable. Define*

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

We would like to find cdf of Y. Note that $\mathbb{P}(Y = 1) = \mathbb{P}(X \ge 0)$ and $\mathbb{P}(Y = -1) = \mathbb{P}(X < 0)$. Hence we get

$$\mathbb{P}(Y \le y) = \begin{cases} 0, & y < -1 \\ \mathbb{P}(X < 0), & -1 \le y < 1 \\ 1, & y \ge 1. \end{cases}$$

Theorem 2.2. *Let X be a continuous random variable having pdf f_X. Suppose that $g(x)$ is strictly monotone, differentiable function of x. Then $Y = g(X)$ has a pdf given by*

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))\Big|\frac{d}{dy}g^{-1}(y)\Big| & \text{if } y = g(x) \text{ for some } x \\ 0, & \text{if } y \ne g(x) \text{ for all } x. \end{cases}$$

Proof. Suppose g is increasing. Let $y = g(x)$ for some x. Then

$$F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}\big(X \le g^{-1}(y)\big) = F_X(g^{-1}(y))$$

Differentiating the above equality, we have

$$f_Y(y) = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) = f_X(g^{-1}(y))\Big|\frac{d}{dy}g^{-1}(y)\Big|$$

where in the last equality, we have used the fact that derivative of $g^{-1}(y)$ is non-negative (as $g^{-1}(y)$ is non-decreasing). When $y \ne g(x)$ for all x, then $F_Y(y)$ is either 0 or 1and hence $f_Y(y) = 0$. This completes the proof. ☐

Corollary 2.3. *Let g be piecewise strictly monotone and continuously differentiable i.e., there exist intervals $I_1, I_2, \ldots, I_n$ which partition $\mathbb{R}$ such that g is strictly monotone and continuously differentiable on the interior of each I_i. Then $Y = g(X)$ has a pdf given by*

$$F_Y(y) = \sum_{k=1}^{n} f_X(g_k^{-1}(y))\Big|\frac{d}{dy}g_k^{-1}(y)\Big|$$

where g_k^{-1} is the inverse of g on I_k.

Example 2.9. *Find the pdf of $Y = X^3$ for any continuous nonnegative random variable X with density function f_X.*

Solution: *Let $g(x) = x^3$. Then g is monotone, $g^{-1}(y) = y^{\frac{1}{3}}$ and $\frac{d}{dy}g^{-1}(y) = \frac{1}{3}y^{-\frac{2}{3}}$. By applying Theorem 2.2, we have*

$$f_Y(y) = \frac{1}{3}y^{-\frac{2}{3}}f_X(y^{\frac{1}{3}}).$$

2.3. **Expectation of a function of random variable:**

Definition 2.7 (Expected value of a random variable). Let X be discrete random variable with values $E = \{x_1, x_2, \dots\}$. We define the expected value of X as

$$\mathbb{E}(X) := \sum_i x_i \mathbb{P}(X = x_i), \quad \text{provided } \sum_i |x_i| \mathbb{P}(X = x_i) < +\infty.$$

Let X be a continuous random variable with pdf f_X. We define the expected value of X as

$$\mathbb{E}[X] = \int_{\mathbb{R}} x f_X(x)\, dx, \quad \text{provided } \int_{\mathbb{R}} |x| f_X(x)\, dx < +\infty.$$

Example 2.10. *Let X be a discrete random variable with pmf given by*

$$p_X(x) = \begin{cases} e^{-5}\frac{5^x}{x!}, & x = 0, 1, 2, \dots \\ 0, & \textit{otherwise} \end{cases}$$

Then

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} k e^{-5} \frac{5^k}{k!} = 5e^{-5} \sum_{i=0}^{\infty} \frac{5^i}{i!} = 5.$$

Example 2.11. *Let X be a random variable with pdf*

$$f_X(x) = \begin{cases} \frac{x^{-\frac{1}{2}}}{4\pi}, & 0 \le x \le 4\pi^2 \\ 0, & \textit{otherwise} \end{cases}$$

Then

$$\mathbb{E}[X] = \int_0^{4\pi^2} \frac{x^{\frac{1}{2}}}{4\pi}\, dx = \frac{4}{3}\pi^2.$$

Example 2.12. *Let X be a random variable with pdf given by $f_X(x) = \frac{\alpha}{\alpha^2 + x^2}$ where $\alpha > 0$ is a constant. Observe that*

$$\int_{\mathbb{R}} |x| f_X(x)\, dx = \frac{2\alpha}{\pi} \int_0^{\infty} \frac{x}{\alpha^2 + x^2}\, dx = \frac{\alpha}{\pi} \int_{\alpha^2}^{\infty} \frac{1}{z}\, dz = \infty \; (\because \lim_{M \to \infty} \int_{\alpha^2}^{M} \frac{1}{z}\, dz = \lim_{M \to \infty} \big[\log(M) - \log(\alpha^2)\big])$$

Thus, $\mathbb{E}(X)$ does not exist.

Let X be a given random variable defined on a given probability space and $g : \mathbb{R} \to \mathbb{R}$ be a Borel measurable function. We want to find the expectation of $Y = g(X)$.

Theorem 2.4. *Let X be a random variable on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then*

$$\mathbb{E}[g(X)] = \begin{cases} \sum_x g(x) p_X(x), & \textit{if } X \textit{ is a discrete random variable} \\ \int_{-\infty}^{\infty} g(x) f_X(x)\, dx, & \textit{if } X \textit{ is a continuous random variable} \end{cases}$$

provided the above sum or integral converges absolutely.

Proof. Let X be a discrete random variable that takes the values $x_1, x_2, \dots$. Then $Y := g(X)$ takes the values $g(x_1), g(x_2), \dots$ but some of these values be same. Let y_j, $j \ge 1$ be the distinct values of $g(x_i)$. Then, by grouping, we have

$$\begin{aligned} \sum_i g(x_i) p_X(x_i) &= \sum_j \sum_{i: g(x_i) = y_j} g(x_i) p_X(x_i) = \sum_j y_j \sum_{i: g(x_i) = y_j} p_X(x_i) = \sum_j y_j \mathbb{P}(g(X) = y_j) \\ &= \sum_j y_j \mathbb{P}(Y = y_j) = \mathbb{E}[Y] = \mathbb{E}[g(X)]. \end{aligned}$$

To prove second part, let us first prove the following: for any non-negative random variable Y, one has

$$\mathbb{E}[Y] = \int_0^\infty \mathbb{P}(Y > y)\, dy.$$

We prove this claim for continuous random variable Y with density function f_Y. Indeed,

$$\begin{aligned}\int_0^\infty \mathbb{P}(Y > y)\, dy &= \int_0^\infty \Big(\int_y^\infty f_Y(x)\, dx \Big)\, dy = \int_0^\infty \Big(\int_0^x f_Y(x)\, dy \Big)\, dx \\ &= \int_0^\infty x f_Y(x)\, dx = \mathbb{E}[Y]\,.\end{aligned}$$

For continuous random variable X, let g be non-negative Borel measurable function. Then, we have

$$\begin{aligned}\mathbb{E}[g(X)] &= \int_0^\infty \mathbb{P}(g(X) > y)\, dy = \int_0^\infty \Big(\int_{x: g(x) > y} f_X(x)\, dx \Big)\, dy \\ &= \int_{x: g(x) > 0} \Big(\int_0^{g(x)} f_X(x)\, dy \Big)\, dx = \int_{x: g(x) > 0} g(x) f_X(x)\, dx\,.\end{aligned}$$

□

Example 2.13. *A product that is sold seasonally yields a net profit of b dollars for each unit sold and a net loss of ℓ dollars for each unit left unsold when the season ends. The number of units of the product that are ordered at a specific department store during any season is a random variable having probability mass function $p_i, i \ge 0$. If the store must stock this product in advance, determine the number of units the store should stock so as to maximize its expected profit.*

Solution: *Let X denotes the number of ordered units and S is the stock. Then the profit is defined as*

$$P(S) = \begin{cases} bX - (S - X)\ell, & X \le S \\ Sb, & X > S. \end{cases}$$

Hence the expected profit

$$\begin{aligned}\mathbb{E}(P(S)) &= \sum_{i=0}^{S} [bi - (S - i)\ell] p_i + \sum_{i=S+1}^{\infty} S b p_i = (b + \ell) \sum_{i=0}^{S} i p_i + Sb - (b + \ell) S \sum_{i=0}^{S} p_i \\ &= Sb + (b + \ell) \sum_{i=0}^{S} (i - S) p_i\,. \\ &\implies \mathbb{E}(P(S + 1)) = b(S + 1) + (b + \ell) \sum_{i=0}^{S} (i - S - 1) p_i\,.\end{aligned}$$

Thus,

$$\mathbb{E}(P(S + 1)) - \mathbb{E}(P(S)) = b - (b + \ell) \sum_{i=0}^{S} p_i.$$

Hence, stocking $S + 1$ units will be better than stocking S units whenever

$$\sum_{i=0}^{S} p_i < \frac{b}{b + \ell}. \tag{2.1}$$

Observe that l.h.s of (2.1) is increasing in S where as r.h.s of (2.1) is constant, and hence the inequality will be satisfied for all values of $S \le S^*$ where S^* is the largest value of S satisfying (2.1). Thus, stocking $S^* + 1$ items will lead to a maximum expected profit.

Example 2.14. *Let X be a random variable with pdf given by*

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty$$

where $\sigma > 0$ and $\mu \in \mathbb{R}$ are given constants. We would like to find $\mathbb{E}(X)$ and $\mathbb{E}(X^2)$.

$$\begin{aligned}
\mathbb{E}[X] &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} x e^{-\frac{(x-\mu)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma y + \mu)e^{-\frac{y^2}{2}}\,dy \ (\textit{taking } y = \frac{x-\mu}{\sigma}) \\
&= \mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy + \underbrace{\frac{\sigma}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\,dy}_{=0} = \mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy = \mu,
\end{aligned}$$

where in the last equality, we have used that

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy = \sqrt{2\pi}.$$

$$\begin{aligned}
\mathbb{E}[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma y + \mu)e^{-\frac{y^2}{2}}\,dy \\
&= \frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}}\,dy + \underbrace{\frac{2\sigma\mu}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y e^{-\frac{y^2}{2}}\,dy}_{=0} + \mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy \\
&= -\frac{\sigma^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} y\frac{d}{dy}e^{-\frac{y^2}{2}} + \mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy \\
&= (\mu^2 + \sigma^2)\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,dy - \frac{\sigma^2}{\sqrt{2\pi}}(ye^{-\frac{y^2}{2}})\Big|_{-\infty}^{\infty} = \sigma^2 + \mu^2.
\end{aligned}$$

We now state some important propeties of the expected value of a random variable, whose proof is simple and hence left as excercise.

Theorem 2.5. *Let X be a random variable defined on a given probability space.*

i) If X is non-negative, i.e., $\mathbb{P}(X \ge 0) = 1$ and $\mathbb{E}(X)$ exists, then $\mathbb{E}(X) \ge 0$.
ii) For any constant c, $\mathbb{E}[c] = c$.
iii) If X is bounded i.e., there exists $M > 0$ such that $\mathbb{P}(|X| > M) = 0$, then $\mathbb{E}[X]$ exists.
iv) Let g and h be two functions such that $g(X)$ and $h(X)$ are random variables and both $\mathbb{E}[h(X)]$ and $\mathbb{E}[g(X)]$ exist. Then for any $\alpha, \beta \in \mathbb{R}$, expectation of the random variable $\alpha g(X) + \beta h(X)$ exists and given by

$$\mathbb{E}[\alpha g(X) + \beta h(X)] = \alpha\mathbb{E}[g(X)] + \beta\mathbb{E}[h(X)].$$

In other words, $\mathbb{E}[\cdot]$ is linear. Moreover, if $g(x) \le h(x)$ for all x, then

$$\mathbb{E}[g(X)] \le \mathbb{E}[h(X)].$$

In particular, one has the following:

$$\big|\mathbb{E}[X]\big| \le \mathbb{E}[|X|].$$

If $g(x) = x^n$, then $\mathbb{E}[g(X)] = \mathbb{E}[X^n]$ is called *moment* of order n.

Lemma 2.6. *Let X be a random variable such that $\mathbb{E}[X^r]$ exists. Then $\mathbb{E}[X^s]$ exists for all $s < r$.*

Proof. We prove it for continuous random variable. For discrete case, it is similar, and left as excercise. Let X be a continuous random variable with pdf f_X. Let $s < r$. Then for any $x \in \mathbb{R}$, we have $|x^s| < 1 + |x^r|$, and hence

$$\int_{-\infty}^{\infty} |x^s| f_X(x)\,dx < \int_{-\infty}^{\infty} (1 + |x^r|) f_X(x)\,dx = 1 + \int_{-\infty}^{\infty} |x^r| f_X(x)\,dx < +\infty.$$

This shows that $\mathbb{E}[X^s]$ exists for all $s < r$ $\square$

We denote by μ the mean of a random variable i.e., $\mu = \mathbb{E}[X]$. It is reasonable to think how far X is from its mean. Instead, we are interested in the quantity $\mathbb{E}[(X-\mu)^2]$.

Definition 2.8 (Variance of a random variable:). Let X be a random variable with finite second moment i.e., $\mathbb{E}[X^2] < \infty$ and mean μ. Then the variance of X, denoted by $\mathrm{Var}(X)$, is defined as

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

We now discuss certain properties of variance.

Theorem 2.7. *The followings hold:*

i) $\mathrm{Var}(X) \geq 0$ *and* $\mathrm{Var}(\alpha) = 0$ *for any constant* α.
ii) $\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}[X])^2$.
iii) $\mathrm{Var}(aX+b) = a^2\mathrm{Var}(X)$.
iv) $\mathrm{Var}(X) = 0$ *if and only if* $\mathbb{P}(X = \mathbb{E}[X]) = 1$.

Proof. Proof of i) : Since $(X-\mu)^2 \geq 0$, we have $\mathbb{E}[(X-\mu)^2] \geq 0$ i.e., $\mathrm{Var}(X) \geq 0$. Moreover, for any constant α, we have $\mathrm{Var}(\alpha) = \mathbb{E}[(\alpha - \mathbb{E}[\alpha])^2] = \mathbb{E}[0] = 0$.
Proof of ii) : By using linearity of expectation, we see that

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

Proof of iii) : By using linearity of expectation and ii), we have for any $a, b \in \mathbb{R}$

$$\begin{aligned}\mathrm{Var}(aX+b) &= \mathbb{E}[(aX+b)^2] - (\mathbb{E}[aX+b])^2 = \mathbb{E}[a^2X^2 + 2abX + b^2] - (a\mathbb{E}[X] + b)^2 \\ &= a^2\{\mathbb{E}[X^2] - (\mathbb{E}[X])^2\} + 2ab\mathbb{E}[X] + b^2 - 2ab\mathbb{E}[X] - b^2 = a^2\mathrm{Var}(X).\end{aligned}$$

Proof of iv) : If $\mathbb{P}(X = \mathbb{E}[X]) = 1$, then it is clear that $\mathrm{Var}(X) = 0$. Suppose $\mathrm{Var}(X) = 0$. If possible, let $\mathbb{P}(X = \mu) < 1$. Then there exists a constant $c > 0$ such that

$$\mathbb{P}((X-\mu)^2 > c) > 0.$$

We then have

$$\begin{aligned}0 = \mathrm{Var}(X) = \mathbb{E}[(X-\mu)^2] &= \mathbb{E}\Big[(X-\mu)^2\mathbf{I}_{\{\omega:(X-\mu)^2>c\}} + (X-\mu)^2\mathbf{I}_{\{\omega:(X-\mu)^2\leq c\}}\Big] \\ &\geq \mathbb{E}\big[(X-\mu)^2\mathbf{I}_{\{\omega:(X-\mu)^2>c\}}\big] > c\mathbb{E}\big[\mathbf{I}_{\{\omega:(X-\mu)^2>c\}}\big] = c\mathbb{P}((X-\mu)^2 > c) > 0 \Rightarrow\Leftarrow!\end{aligned}$$

$\square$

One can consider the central moment of order k, i.e., $\mathbb{E}[(X-\mu)^k]$.

2.4. **Generating functions and their applications:** We will discuss now various type of generating functions and their applications. One of them is probability generating function, which plays an important role in applied probability, in particular in stochastic process.

Definition 2.9 (Probability generating function (pgf)). Let X be a non-negative integer-valued random variable and let $p_i = \mathbb{P}(X = i),\ i = 0, 1, 2, \dots$ with $\sum_{i=0}^{\infty} p_i = 1$. The probability generating function of X is defined by

$$G_X(s) = \mathbb{E}[s^X] = \sum_{i=0}^{\infty} p_i s^i, \quad |s| < 1.$$

Note that, for $s = 1$, $G_X(1) = \mathbb{E}[1] = 1$.

Example 2.15. i): *If $\mathbb{P}(X = c) = 1$, then $G_X(s) = \mathbb{E}[s^X] = s^c$.*
ii): *If $\mathbb{P}(X = 1) = p$ and $\mathbb{P}(X = 0) = 1 - p$ for some $p \in (0, 1)$, then $G_X(s) = (1 - p) + ps$.*
iii): *Let X be a discrete random variable with pmf given by $p_X(k) = e^{-\lambda}\frac{\lambda^k}{k!}$ for $k = 0, 1, 2, \dots$. Then the pgf of X is given by*

$$G_X(s) = \sum_{k=0}^{\infty} s^k e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(s\lambda)^k}{k!} = e^{\lambda(s-1)}.$$

Next theorem is about the application of pgf.

Theorem 2.8. *Let X be a non-negative integer valued random variable with pgf $G_X(\cdot)$. Then*

$$\mathbb{E}[X(X-1)\dots(X-k+1)] = G_X^{(k)}(1)$$

where $G_X^{(k)}(1) = \lim_{s\to 1} G_X^{(k)}(s)$.

Proof. Note that, since the series $\sum_{i=0}^{\infty} p_i s^i$ converges in $|s| < 1$, it converges uniformly and therefore term by term differentiation is possible at points s satisfying $|s| < 1$. Thus, we have

$$G_X^{(k)}(s) = \sum_i s^{i-k} i(i-1)\dots(i-k+1)p_i = \mathbb{E}[s^{X-k}X(X-1)\dots(X-k+1)].$$

This sum is convergent for $|s| < 1$. To prove the theorem, we will use the following theorem, called **Abel's theorem**. It states that if $a_i \ge 0$ for all i, and $G_a(s) = \sum_{i=0}^{\infty} a_i s^i$ is finite for $|s| < 1$, then

$$\lim_{s\to 1} G_a(s) = \sum_{i=0}^{\infty} a_i$$

whether the sum is finite or equals to $+\infty$. Using Abel's theorem, we have

$$\lim_{s\to 1} G_X^{(k)}(s) = \mathbb{E}[X(X-1)\dots(X-k+1)] = \sum_i i(i-1)\dots(i-k+1)p_i.$$

This completes the proof. □

Suppose that $\mathbb{E}[X^2] < +\infty$. Then we can calculate $\mathrm{Var}(X)$ in terms of pgf. We have

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \mathbb{E}[X(X-1)] + \mathbb{E}[X] - (\mathbb{E}[X])^2 = G_X^{(2)}(1) + G_X^{(1)}(1) - \{G_X^{(1)}(1)\}^2.$$

One can easily check that

$$p_i = \mathbb{P}(X = i) = \frac{1}{n!}\frac{d^n}{ds^n}G_X(s)\Big|_{s=0}.$$

Probability generating function uniquely determine the distribution function. In particular, if two random variable have the same pgf in some interval containing zero, then the random variables have same distribution.

Another important generating function is called *moment generating function*.

Definition 2.10 (**Moment generating function (mgf):**)**.** For a given random variable X, we define its moment generating function, denoted by $m_X(\cdot)$, as

$$m_X(t) = \mathbb{E}[e^{tX}]$$

provided $\mathbb{E}[e^{tX}]$ exists in a small nbd. of the origin.

Example 2.16. *Let X be a discrete random variable with pmf $p_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$. Then*

$$m_X(t) = e^{-\lambda} \sum_{i=0}^{\infty} e^{tk} \frac{\lambda^k}{k!} = e^{-\lambda(1-e^t)} \quad \forall t.$$

Example 2.17. *Let X be a random variable with pdf given by*

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & x \le 0. \end{cases}$$

Then mgf of X is given by

$$m_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} 2e^{-2x}\, dx = 2\int_0^\infty e^{-(2-t)x}\, dx = \frac{2}{2-t}, \quad \text{for } t < 2.$$

Example 2.18. *Let X be a discrete random variable with pmf $p_X(k) = \frac{6}{\pi^2}\frac{1}{k^2}$ for $k = 0, 1, 2, \ldots$. Then for any $s > 0$, $\frac{1}{\pi^2}\sum_{k=1}^\infty \frac{e^{sk}}{k^2}$ is infinite and hence $m_x(\cdot)$ does not exist. In fact, one can check that $\mathbb{E}[X] = \infty$.*

Theorem 2.9. *Let X be a random variable whose moment generating function $m_X(\cdot)$ exists. Then for any $r \in \mathbb{N}$, the r-th moment $\mathbb{E}[X^r]$ exists. Moreover, $\mathbb{E}[X^r]$ can be calculated as follows:*

$$\frac{d^r}{dx^r} m_X(s)\Big|_{s=0} = \mathbb{E}[X^r], \quad r \ge 1.$$

Proof. Since $m_X(t)$ exists, there exists $a > 0$ such that $m_X(a)$ and $m_X(-a)$ are finite for some $a > 0$. We claim that for any $s \in (-a, a)$, $m_X(s)$ exists. Indeed,

$$\mathbb{E}[e^{sX}] = \mathbb{E}\big[e^{sX}\big(\mathbf{1}_{X>0} + \mathbf{1}_{X\le 0}\big)\big] \le \mathbb{E}[e^{aX} + e^{-aX}] = m_X(a) + m_X(-a) < \infty\,.$$

Now, for any $k \in \mathbb{R}^*$, there exists $s \in (-a, a)$ such that $0 < sk < a$. Therefore, since $s|X| \le e^{s|X|}$ we have

$$\begin{aligned} & s^k|X|^k \le e^{sk|X|} \le e^{aX} + e^{-aX} \\ & \implies \mathbb{E}[|X|^k] \le s^{-k}\mathbb{E}[e^{aX} + e^{-aX}] = s^{-k}\big(m_X(a) + m_X(-a)\big) < \infty\,. \end{aligned}$$

Observe that

$$\mathbb{E}[\sum_{i=0}^n \frac{t^i}{i!} X^i] = \sum_{i=0}^n \frac{t^i}{i!}\mathbb{E}[X^i] \le \big(m_X(a) + m_X(-a)\big) \sum_{i=0}^n \frac{(\frac{t}{s})^i}{i!}$$

and hence we see that

$$m_X(t) = \mathbb{E}[e^{tX}] = \mathbb{E}[\sum_{i=0}^\infty \frac{t^i}{i!} X^i] = \sum_{i=0}^\infty \frac{t^i}{i!}\mathbb{E}[X^i]\,.$$

It is a power series of $m_X(t)$ around $t = 0$, and hence term by term differentiation is possible for $|t| < a$. This implies that

$$\frac{d^r}{ds^r} m_X(s) = \sum_{i=0}^\infty \frac{s^i}{i!}\mathbb{E}[X^{r+i}] \implies \frac{d^r}{dx^r} m_X(s)\Big|_{s=0} = \mathbb{E}[X^r], \quad r \ge 1.$$

□

Next we state an important theorem regarding mgf without proof.

Theorem 2.10. *If $m_X(t)$ exists, then it is unique. Moreover, it determines the distribution uniquely. In particular, if X and Y be two random variables whose moment generating functions exist, and $m_X(t) = m_Y(t)$ for all t, then X and Y have the same distribution.*

We now discuss another important generating function which always exists.

Definition 2.11 (Characteristic function:). For a random variable X, we define its characteristic function $\phi_X : \mathbb{R} \to \mathbb{C}$ via

$$\phi_X(t) = \mathbb{E}[e^{itX}] = \mathbb{E}[\cos(tX)] + i\,\mathbb{E}[\sin(tX)], \quad i = \sqrt{-1}.$$

We now prove some important properties of characteristic function.

Theorem 2.11. *Let $\phi_X(\cdot)$ be the characteristic function of a random variable X. Then the followings hold:*

i) $\phi_X(0) = 1$ *and* $|\phi_X(t)| \le 1$ *for all* $t \in \mathbb{R}$.
ii) $\phi_X(\cdot)$ *is uniformly continuous on* $\mathbb{R}$.
iii) *If* $a, b \in \mathbb{R}$ *and* $Y = aX + b$, *then* $\phi_Y(t) = e^{itb}\phi_X(at)$.
iv) *If* $\mathbb{E}[|X|^m < +\infty]$, *then* ϕ_X *is* m*- times continuously differentiable and*

$$\frac{d^m}{dt^m}\phi_X(t)\Big|_{t=0} = i^m\,\mathbb{E}[X^m].$$

v) *If* X *and* Y *are random variables such that* $\phi_X(t) = \phi_Y(t)$ *for all* $t \in \mathbb{R}$, *then* X *and* Y *have same distribution.*

Proof. Proof of i): Since $\cos(tX)$ and $\sin(tX)$ are bounded functions, expectation exists and hence $\phi_X(t)$ exists for all $t \in \mathbb{R}$. Moreover, $\phi_X(0) = \mathbb{E}[1] = 1$. Furthermore,

$$|\phi_X(t)| \le \mathbb{E}[|e^{itX}|] = \mathbb{E}[1] = 1, \quad \forall\, t \in \mathbb{R}.$$

Proof of ii): Let $t \in \mathbb{R}$. For any h, we have

$$\big|\phi_X(t+h) - \phi_X(t)\big| = \big|\mathbb{E}[e^{itX}(e^{ihX} - 1)]\big| \le \mathbb{E}[|e^{itX}|\,|e^{ihX} - 1|] = \mathbb{E}[|e^{ihX} - 1|].$$

Set $Y(h) = |e^{ihX} - 1|$. Then $Y(h) \to 0$ as $h \to 0$ and $|Y_h| \le 2$. Hence by bounded convergence theorem, $\mathbb{E}[Y(h)] \to 0$. Thus, we conclude that $\phi_X(\cdot)$ is uniformly continuous on $\mathbb{R}$.
Proof of iii): Observe that Y is a random variable. Moreover, by using linearity of expectation, we have

$$\phi_Y(t) = \mathbb{E}[e^{it(aX+b)}] = \mathbb{E}[e^{itb}e^{ita\,X}] = e^{itb}\mathbb{E}[e^{i(at)X}] = e^{itb}\phi_X(at).$$

Proof of iv): We prove the result for $m = 1$. The general case can be established analogously by recurrance. Fix $t \in \mathbb{R}$. For any $h \in \mathbb{R}$, we have, by using linearity of expectation,

$$\frac{1}{h}[\phi_X(t+h) - \phi_X(t)] = \frac{1}{h}\mathbb{E}[e^{itX}(e^{ihX} - 1)] = \mathbb{E}[e^{itX} \cdot \frac{e^{ihX} - 1}{h}]$$

Note that $\frac{e^{ihx}-1}{h} \to ix$ as $h \to 0$ and $|\frac{e^{ihx}-1}{h}| \le 2|x|$ for small h. Since $\mathbb{E}[|X|] < +\infty$, by dominated convergence theorem, we have

$$\begin{aligned}\lim_{h\to 0}\frac{1}{h}[\phi_X(t+h) - \phi_X(t)] &= \lim_{h\to 0}\mathbb{E}[e^{itX} \cdot \frac{e^{ihX} - 1}{h}] = \mathbb{E}\big[e^{itX}\lim_{h\to 0}\frac{e^{ihX} - 1}{h}\big] \\ &= \mathbb{E}\big[e^{itX} iX\big] = i\,\mathbb{E}[Xe^{itX}] < +\infty\end{aligned}$$

So, $\phi'_X(t)$ exists and $\phi'_X(t) = i\,\mathbb{E}[Xe^{itX}]$. We now show that $\phi'_X(\cdot)$ is continuous. Norice that

$$\big|\phi'_X(t+h) - \phi'_X(t)\big| = \big|i\,\mathbb{E}[Xe^{itX}(e^{ihX} - 1)]\big| \le \mathbb{E}[|X|Y(h)]$$

where $Y(h) := |e^{ihX} - 1|$. Notice that $|X|Y(h) \to 0$ as $h \to 0$ and $\big||X|Y(h)\big| \le 2|X|$. Since $\mathbb{E}[|X|] < +\infty$, by dominated convergence theorem, we have

$$\lim_{h\to 0}\mathbb{E}[|X|Y(h)] = \mathbb{E}[\lim_{h\to 0}|X|Y(h)] = 0.$$

Hence we conclude that $\phi_X'(\cdot)$ is uniformly continuous on $\mathbb{R}$. $\square$

Example 2.19. *Let X be a random variable with pdf $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$. Then for any $t \in \mathbb{R}$, we have*

$$\begin{aligned}\phi_X(t) = \mathbb{E}[e^{itX}] = \int_0^\infty \lambda e^{-\lambda x} e^{itx}\,dx = \lambda\Big\{\int_0^\infty \cos(tx)e^{-\lambda x}\,dx + i\int_0^\infty \sin(tx)e^{-\lambda x}\,dx\Big\}\\ = (\lambda + i\,t)\int_0^\infty \cos(tx)e^{-\lambda x}\,dx := (\lambda+it)I(t).\end{aligned}$$

We now calculate $I(t)$. By applying integration by parts formula, we have

$$\begin{aligned}I(t) = -\frac{1}{\lambda}\int_0^\infty \cos(tx)\frac{d}{dx}e^{-\lambda x} = -\frac{t}{\lambda}\int_0^\infty \sin(tx)e^{-\lambda x} + \frac{1}{\lambda} = -\frac{t^2}{\lambda^2}I(t) + \frac{1}{\lambda}\\ \implies I(t) = \frac{\lambda}{t^2+\lambda^2}\,.\end{aligned}$$

Thus, we have

$$\phi_X(t) = \frac{\lambda(\lambda+it)}{t^2+\lambda^2} = \frac{\lambda}{\lambda - it}.$$

Example 2.20. *Let X be a continuous random variable with density function f_X given by*

$$f_X(x) = \begin{cases}1, & 0<x<1\\ 0, & \textit{otherwise.}\end{cases}$$

then for any $t \neq 0$, we have

$$\phi_X(t) = \int_0^1 \cos(tx)\,dx + i\int_0^1 \sin(tx)\,dx = \frac{1}{t}[\sin(t) - i\cos(t) + i] = \frac{1}{it}(e^{it}-1).$$