Chapter 2: Random Variables

1 Definition of Random Variable

- When an experiment is performed, we are often interested in some numerical characteristic of the result.

- Here are some practical examples.

  - An experiment is to examine 100 electronic components, our interest is “the number of defectives".
  - Flipping a coin 100 times, our interest could be the number of heads obtained, instead of the “H" and “T" sequence.

- This motivates us to assign a numerical value to a possible outcome of an experiment.

Definition 1 (Random Variable)
*Let S be sample space for an experiment. A **function** X, which assigns a real number to every $s \in S$ is called a **random variable**.*

- So random variable X is a function from S to $\mathbb{R}$:

$$X : S \mapsto \mathbb{R}.$$

- For convenience, hereafter, we simplify "**random variable**" as "**RV**".

Example 2.1

- Let $S = \{HH, HT, TH, TT\}$ be a sample space associated with the experiment of flipping two coins.
- Define the RV:

$$X = \text{number of heads obtained.}$$

- Note that X is a **function** from S to $\mathbb{R}$, the set of real numbers:

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

  The range of X is $R_X = \{0, 1, 2\}$.

REMARK

- We use upper case letters $X, Y, Z, X_1, X_2, \ldots$ to denote **random variables**.
- We use lower case letters x, y, z, x_1, x_2 to denote their **observed values** in the experiment.
- The set $\{X = x\}$ is a subset of S, in the sense:

$$\{X = x\} = \{s \in S : X(s) = x\}.$$

- Likewise, the set $\{X \in A\}$, for A being a subset of $\mathbb{R}$, is also a subset of S:
$$\{s \in S : X(s) \in A\}.$$

- This gives $P(X = x)$ and $P(X \in A)$ based on probability defined on S:

$$\begin{aligned} P(X = x) &= P(\{s \in S : X(s) = x\}) \\ P(X \in A) &= P(\{s \in S : X(s) \in A\}) \end{aligned}$$

Example 2.2

- Revisit Example 2.1; $S = \{HH, HT, TH, TT\}$ is the sample space of flipping two coins. $X =$ number of heads obtained.
- Then $\{X = 0\} = \{TT\}$; $\{X = 1\} = \{HT, TH\}$; $\{X = 2\} = \{HH\}$; $\{X \geq 1\} = \{HT, TH, HH\}$.
- $P(X = 0) = P(TT) = 1/4$; $P(X = 1) = P(\{HT, TH\}) = 2/4$; $P(X = 2) = P(HH) = 1/4$; $P(X \geq 1) = P(\{HT, TH, HH\}) = 3/4$.

- We can summarize the probabilities of the RV X as a table:

x	0	1	2
$P(X = x)$	1/4	1/2	1/4

2 Probability Distributions

- We introduce the distribution of two types of RVs: **discrete** and **continuous**.
- In particular, denote by X the RV, and its range by R_X.
  - **Discrete**: the number of values in R_X is **finite** or **countable**; that is we can write $R_X = \{x_1, x_2, x_3, \ldots\}$.
  - **Continuous**: R_X is an **interval** or a **collection of intervals**.

Discrete Probability Distributions

- For a discrete RV X, we can always write $R_X = \{x_1, x_2, x_3, \ldots\}$.
- Each $x_i \in R_X$, there is a probability that X takes this value, i.e., $P(X = x_i)$.
- We can define a function $f(x) = P(X = x)$.

  Note that $f(x_i) = P(X = x_i)$ for $x_i \in R_X$, and $f(x) = 0$ for $x \notin R_X$.
- $f(x)$ is called the **probability function, p.f.** (or **probability mass function, p.m.f.**) of X.
- The collection of pairs $(x_i, f(x_i)), i = 1, 2, 3, \ldots,$ is called the **probability distribution** of X.

The p.f. $f(x)$ of a discrete RV **must** satisfy:

(1) $f(x_i) \geq 0$ for all $x_i \in R_X$;

(2) $f(x) = 0$ for all $x \notin R_X$;

(3) $\sum_{i=1}^{\infty} f(x_i) = 1$, or $\sum_{x_i \in R_X} f(x_i) = 1$.

For any set $B \subset \mathbb{R}$, we have

$$P(X \in B) = \sum_{x_i \in B \cap R_X} f(x_i).$$

Example 2.3

- Revisit Examples 2.1 and 2.2. RV X is the number of heads when flipping two coins.
- The p.f. of X is given below

x	0	1	2
$f(x)$	1/4	1/2	1/4

- $f(x)$ satisfies (1) $f(x_i) \geq 0$ for $x_i = 0, 1$, or 2; (2) $f(x) = 0$ for other x; (3) $f(0) + f(1) + f(2) = 1$.
- $B = [1, \infty)$; then $P(X \in B) = f(1) + f(2) = 3/4$.

Continuous Probability Distributions

- For a continuous RV X, R_X is an interval or a collection of intervals.

- For any $x \in \mathbb{R}$, we must have $P(X = x) = 0$.

- The **probability function, p.f.**, (or **probability density function, p.d.f.**) is defined to quantify the probability that X is in a certain range.

The **p.d.f.** of a continuous RV X, denoted by $f(x)$, is a function that satisfies:

(1) $f(x) \geq 0$ for all $x \in R_X$; and $f(x) = 0$ for $x \notin R_X$.

(2) $\displaystyle\int_{R_X} f(x)dx = 1$.

(3) For any a and b such that $a \leq b$,

$$P(a \leq X \leq b) = \int_a^b f(x)dx.$$

Note: (2) is equivalent to $\displaystyle\int_{-\infty}^{\infty} f(x)dx = 1$, since $f(x) = 0$ for $x \notin R_X$.

Remark

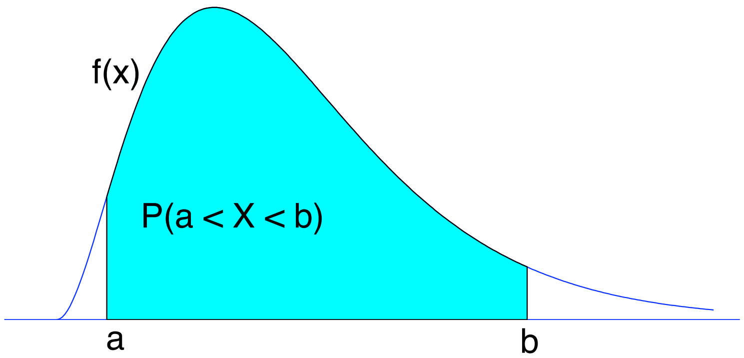
- For any arbitrary specific value x_0, we have

$$P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0.$$

  This gives an example of "$P(A) = 0$, but A is not necessarily $\emptyset$."
  Furthermore, we have

$$P(a < X < b) = P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b) = \int_a^b f(x)\,\mathrm{dx}\,.$$

- They all represent the area under the graph of $f(x)$ between $x = a$ and $x = b$.



- To check that a function $f(x)$ is a p.d.f., it suffices to check (1) and (2), namely,

  (1) $f(x) \geq 0$ for all $x \in R_X$; and $f(x) = 0$ for $x \notin R_X$.

  (2) $\displaystyle\int_{R_X} f(x)dx = 1$.

Example 2.4 Let X be a continuous RV with p.d.f. given by

$$f(x) = \begin{cases} cx, & \text{for } 0 < x < 1; \\ 0, & \text{otherwise.} \end{cases}$$

(a) Find the value of c;

(b) Find $P(X \leq 1/2)$.

Solution:

(a) Since

$$\int_{-\infty}^{\infty} f(x)dx = \int_0^1 cxdx = c \cdot \frac{x^2}{2}\bigg|_0^1 = c/2,$$

we set $c/2 = 1$, and result in $c = 2$.

(b)

$$P(X \leq 1/2) = \int_{-\infty}^{1/2} f(x)dx = \int_0^{1/2} 2xdx = 1/4.$$

3 Cumulative Distribution Function

Definition 2
For any RV X, we define its cumulative distribution function (c.d.f.) by

$$F(x) = P(X \leq x).$$

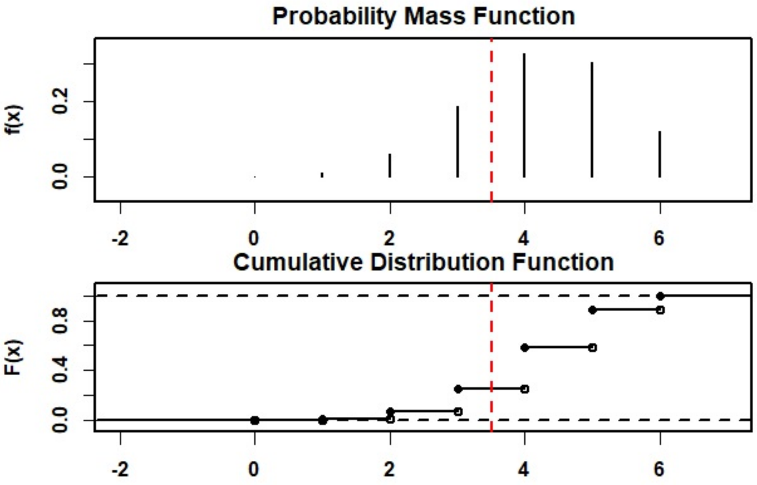
Note: This definition is applicable for X to be either a discrete or a continuous RV.

c.d.f. for Discrete RV

- If X is a **discrete RV**, we have

$$\begin{aligned} F(x) &= \sum_{t \in R_X; t \leq x} f(t) \\ &= \sum_{t \in R_X; t \leq x} P(X = t) \end{aligned}$$

- The c.d.f. of a discrete RV is a step function.



- For any two numbers $a < b$, we have

$$P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - F(a-),$$

where "$a-$" represents the largest value in R_X, that is $< a$. More mathematically,

$$F(a-) = \lim_{x \uparrow a} F(x).$$

Example 2.5

- Revisit Examples 2.1 and 2.2. RV X is the number of heads of flipping two fair coins, it has the p.f.:

x	0	1	2
$f(x)$	1/4	1/2	1/4

- We have $F(0) = f(0) = 1/4$; $F(1) = f(0) + f(1) = 3/4$; $F(2) = f(0) + f(1) + f(2) = 1$.

- We therefore obtain the c.d.f.:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$

Example 2.6 Take the c.d.f. derived from Example 2.5:

$$F(x) = \begin{cases} 0, & x < 0 \\ 1/4, & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}.$$

Derive the corresponding p.f. (pretending that the c.d.f. is only information available for this distribution).

Solution:

- As $F(\cdot)$ only has four possible values, so the distribution is a discrete distribution.

- We obtain $R_X = \{0, 1, 2\}$, which are the jumping points of $F(\cdot)$. It is also the set so that $f(x)$ is non-zero.

- We have

$$
\begin{aligned}
f(0) &= P(X=0) = F(0) - F(0-) = 1/4 - 0 = 1/4; \\
f(1) &= P(X=1) = F(1) - F(1-) = 3/4 - 1/4 = 1/2; \\
f(2) &= P(X=2) = F(2) - F(2-) = 1 - 3/4 = 1/4.
\end{aligned}
$$

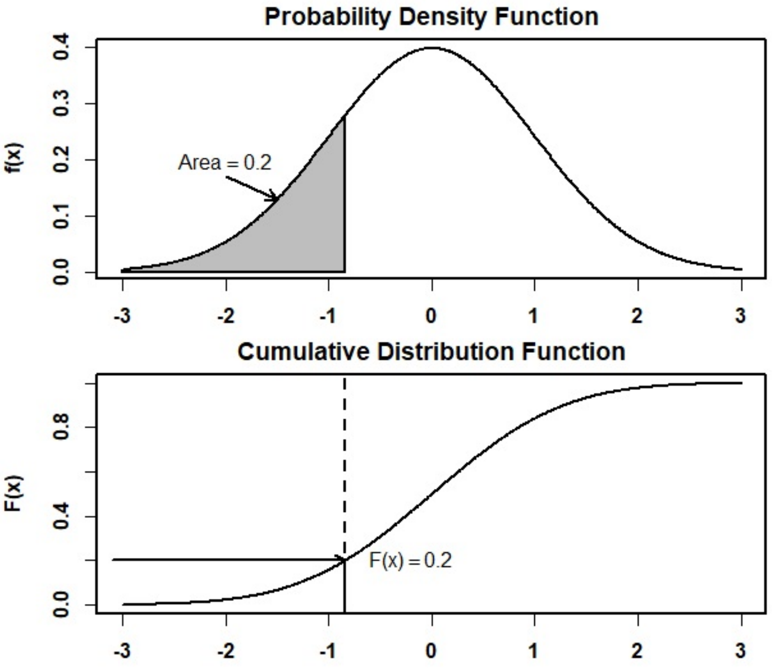
c.d.f. for Continuous RV

- If X is a continuous RV,

$$F(x) = \int_{-\infty}^{x} f(t)dt.$$

$$f(x) = \frac{dF(x)}{dx}.$$

- $P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a).$



Example 2.7

- The p.d.f. of a RV X is given by

$$f(x) = \begin{cases} 2x & 0 \le x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

- The c.d.f. of X is

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \begin{cases} 0 & x < 0 \\ x^2 & 0 \le x < 1 \\ 1 & 1 \le x \end{cases} \end{aligned}$$

Example 2.8 Take the c.d.f. derived from Example 2.7:

$$F(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

Derive the corresponding p.f. (pretending that the c.d.f. is only information available for this distribution).

Solution:

- $F(x)$ is a c.d.f. for a continuous distribution, because when it is not equal to 0 and 1, it assumes different values in the interval $x \in [0,1)$.

- $f(x) = 0$ when $x \notin [0,1)$ because $\dfrac{d(0)}{dx} = \dfrac{d(1)}{dx} = 0$.

- $f(x) = \dfrac{d(x^2)}{dx} = 2x$ when $x \in [0,1)$.

Remark

- No matter whether X is discrete or continuous, $F(x)$ is non-decreasing. In the sense that for any $x_1 < x_2$, $F(x_1) \leq F(x_2)$.

- p.f. and c.d.f. have a one-to-one correspondence. That is, with p.f. given, c.d.f. is uniquely determined; and vice versa.

- The ranges of $F(x)$ and $f(x)$ satisfy:
  - $0 \leq F(x) \leq 1$;
  - for discrete distribution, $0 \leq f(x) \leq 1$;
  - for continuous distribution, $f(x) \geq 0$, but **NO NEED** that $f(x) \leq 1$.

4 EXPECTATION AND VARIANCE OF A RV

- For a RV X, one natural practical question is: what is the **average value** of X, if the corresponding experiment is repeated many times.

  For example, X is the number of heads for flipping a coin. We may want to know the average number of heads if we repeat to flip the coin "continuously".

- Such an average, over a long run, is called the "**mean**" or "**expectation**" of X.

Definition 3 (Expectation of discrete RV)
*Let X be a discrete RV with $R_X = \{x_1, x_2, x_3, \ldots.\}$ and p.f. $f(x)$. The "**expectation**" or "**mean**" of X is defined by*

$$E(X) = \sum_{x_i \in R_X} x_i f(x_i).$$

By convention, we also denote $\mu_X = E(X)$.

Definition 4 (Expectation of continuous RV)
*Let X be a continuous RV with p.f. $f(x)$. The "**expectation**" or "**mean**" of X is defined by*

$$\mu_X = E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_{x \in R_X} x f(x) dx.$$

Note: The expected value is not necessarily a possible value of the random variable X.

Example 2.9 Suppose we toss a fair die and the upper face is recorded as X. We have $P(X = k) = 1/6$ for $k = 1, 2, 3, 4, 5, 6$, and

$$E(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = 3.5.$$

Example 2.10 The p.d.f. of weekly gravel sales X is

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

Then we have

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \frac{3}{2}(1-x^2) dx \\ &= \frac{3}{2}\int_0^1 (x - x^3) dx = \frac{3}{2}\left(\frac{x^2}{2} - \frac{x^4}{4}\right)\Bigg|_0^1 = 3/8. \end{aligned}$$

Properties of Expectation

(1) Let X be a random variable, and let a and b be any real numbers,

$$E(aX+b) = aE(X)+b.$$

(2) Let X and Y be two random variables, we have

$$E(X+Y) = E(X)+E(Y).$$

(3) Let $g(\cdot)$ be an arbitrary function.

- If X is a **discrete** RV with p.m.f. $f(x)$ and range R_X,

$$E[g(X)] = \sum_{x \in R_X} g(x) f(x).$$

- If X is a **continuous** RV with p.d.f. $f(x)$ and range R_X,

$$E[g(X)] = \int_{R_X} g(x) f(x) dx.$$

Variance

Let $g(x) = (x - \mu_X)^2$, this gives the definition of the **variance** for X.

Definition 5 (Variance)
Let X be a RV. The ***variance*** *of X is defined by*

$$\sigma_X^2 = V(X) = E(X - \mu_X)^2.$$

Remark

- The definition is applicable no matter whether X is discrete or continuous.

- If X is a **discrete** RV with p.m.f. $f(x)$ and range R_X,

$$V(X) = \sum_{x \in R_X} (x - \mu_X)^2 f(x).$$

- If X is a **continuous** RV with p.d.f. $f(x)$,

$$V(X) = \int_{-\infty}^{\infty} (x - \mu_X)^2 f(x) dx.$$

- For any X, $V(X) \geq 0$, and "$=$" holds if and only $P(X = E(X)) = 1$, or more intuitively, X is a **constant**.
- Let a and b be any real numbers, then $V(aX + b) = a^2 V(X)$.
- The variance can also be computed by an alternative formula:

$$V(X) = E(X^2) - [E(X)]^2.$$

- The positive square root of the variance is defined as the "**standard deviation**" of X:

$$\sigma_X = \sqrt{V(X)}.$$

Example 2.11 Let the p.f. of a RV X be given by

x	-1	0	1	2
$f(x)$	$1/8$	$2/8$	$1/8$	$4/8$

Find $E(X)$ and $V(X)$.
Solution:

$$\begin{aligned} E(X) &= \sum_{x \in R_X} x f(x) \\ &= (-1)\left(\frac{1}{8}\right) + 0\left(\frac{2}{8}\right) + 1\left(\frac{1}{8}\right) + 2\left(\frac{4}{8}\right) = 1. \end{aligned}$$

$$
\begin{aligned}
V(X) &= \sum_{x \in R_X} [x - E(X)]^2 f(x) = \sum_{x \in R_X} [x-1]^2 f(x) \\
&= (-1-1)^2 \left(\frac{1}{8}\right) + (0-1)^2 \left(\frac{2}{8}\right) \\
&\quad + (1-1)^2 \left(\frac{1}{8}\right) + (2-1)^2 \left(\frac{4}{8}\right) = \frac{5}{4}.
\end{aligned}
$$

Example 2.12 Denote by X the amount of time that a book on reserve at the library is checked out by a randomly selected student. Suppose X has the probability density function

$$f(x) = \begin{cases} x/2, & 0 \le x < 2 \\ 0, & \text{otherwise} \end{cases}$$

Compute $E(X), V(X)$, and σ_X.

Solution:

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = \int_0^2 x \cdot x/2 dx = \left. \frac{x^3}{6} \right|_0^2 = 4/3.$$

We use $V(X) = E(X^2) - [E(X)]^2$ to compute $V(X)$,

$$E(X^2) = \int_0^2 x^2 \cdot x/2 dx = \int_0^2 x^3/2 dx = \left. \frac{x^4}{8} \right|_0^2 = 2.$$

$$V(X) = E(X^2) - [E(X)]^2 = 2 - (4/3)^2 = 2/9.$$

$$\sigma_X = \sqrt{V(X)} = \sqrt{2}/3.$$