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
- 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 $R_X = \{0, 1, 2\}$.

L–example 2.1

- A coin is thrown until a "head" occurs.

$$S = \{H, TH, TTH, TTTH, TTTTH, \cdots\}$$

- Let $X =$ the number of "trials" required. We then have

$$X(H) = 1,\ X(TH) = 2,\ X(TTH) = 3,\ \cdots,\ \text{and so on.}$$

- $R_X = \{1, 2, 3, \ldots, \}$

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 3

- Revisit Example 2; $S = \{HH, HT, TH, TT\}$ 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

L–example 2.2

- When a pair of fair dice is rolled, what is the probability that a sum of 3 is obtained?

$$S = \left\{ (x_1, x_2) \middle| x_1 = 1,2,3,4,5,6;\ x_2 = 1,2,3,4,5,6 \right\}.$$

- $X =$ the sum of two dice. That is for any $(x_1, x_2) \in S$,

$$X((x_1, x_2)) = x_1 + x_2.$$

- The range of X is

$$R_X = \{2, 3, 4, \ldots, 12\}.$$

- Since $\{X = 3\} = \{(1,2), (2,1)\}$, we have

$$P(X=3) \quad = \quad P(\{(1,2),(2,1)\}) = 2/36.$$

- The probabilities of other possible values for X can be found similarly, and are tabulated below:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

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 $R_X = \{x_1, x_2, x_3, \dots\}$.
- 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 $(x_i, f(x_i)), i = 1, 2, 3, \dots$, 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 1

- Revisit Examples 2 and 3. 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 $P(X \in B) = f(1) + f(2) = 3/4$.

L–example 2.3

Six lots of components are ready to be shipped by a certain supplier. The number of defective components in each lot is as follows:

Lot	1	2	3	4	5	6
# of defectives	0	2	0	1	2	0

- One of the lots is to be **randomly** selected and shipped to a customer.
- Let $X = \#$ of defectives in the shipped lot.
- Then $R_X = \{0, 1, 2\}$.
- The lots are selected randomly, so each has the same probability to be chosen.
- Let $f(x)$ be the p.f. of X.
- We have
  - $f(0) = P(X = 0) = P(\text{lot 1 or 3 or 6 is selected}) = 3/6.$
  - $f(1) = P(X = 1) = P(\text{lot 4 is selected}) = 1/6.$
  - $f(2) = P(X = 2) = P(\text{lot 2 or 5 is selected}) = 2/6.$
- The probability function of X can be summarized by

x	0	1	2
$f(x)$	$1/2$	$1/6$	$1/3$

- It satisfies all the properties of probability functions.
- If $B = \{0, 2\}$, $P(X \in B) = f(0) + f(2) = 1/2 + 1/3 = 5/6$.

L–example 2.4

(a) Find the constant c, such that

$$f(x) = cx, \quad \text{for } x = 1, 2, 3, 4,$$

and 0 otherwise, is a probability function of a random variable X.

(b) Compute $P(X \geq 3)$.

<u>Solution</u>:

(a) Based on the property $\sum_{i=1}^{\infty} f(x_i) = 1$, we have

$$f(x_1) + f(x_2) + f(x_3) + f(x_4) = 1,$$

which is

$$c + 2c + 3c + 4c = 1.$$

Therefore $c = 1/10$.

(b) $P(X \geq 3) = f(3) + f(4) = 3/10 + 4/10 = 7/10.$

L–example 2.5

- Consider a group of five potential blood donors: A, B, C, D and E, of whom only A and B have type O+ blood.
- Five blood samples, one from each individual , will be typed in random order until an O+ individual is identified.

Solution:

- Let Y = # of typing needed to identify an O+ individual.
- Let O_i and O_i' be the events that an O+ and a non-O+ individual is typed in the ith typing

$$\begin{aligned} f(1) &= P(Y=1) = P(O_1) = 2/5 = 0.4, \\ f(2) &= P(Y=2) = P(O_1' \cap O_2) = P(O_1')P(O_2|O_1') \\ &= \frac{3}{5} \cdot \frac{2}{4} = 0.3, \end{aligned}$$

$$\begin{aligned} f(3) &= P(O_1')P(O_2'|O_1')P(O_3|O_1' \cap O_2') \\ &= \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{2}{3} = 0.2, \\ f(4) &= P(Y=4) \\ &= P(O_1')P(O_2'|O_1')P(O_3'|O_1' \cap O_2')P(O_4|O_1' \cap O_2' \cap O_3') \\ &= \frac{3}{5} \cdot \frac{2}{4} \cdot \frac{1}{3} \cdot \frac{2}{2} = 0.1, \end{aligned}$$

and $f(y) = 0$ if $y \neq 1,2,3,4$.

- Then the probability function of Y is

y	1	2	3	4
$f(y)$	0.4	0.3	0.2	0.1

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:

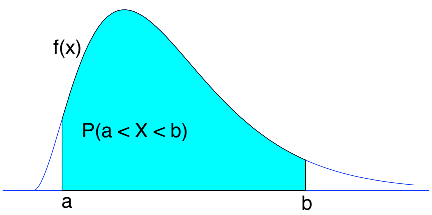
- For any arbitrary specific value x_0, we have

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

  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

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

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

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

(b)

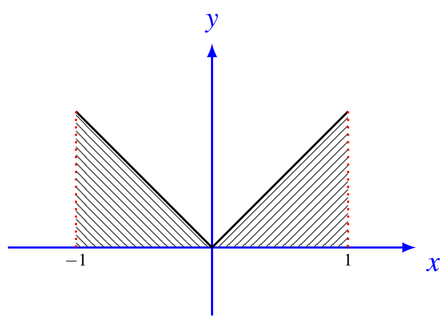
$$P(X \leq 1/2) = \int_{-\infty}^{1/2} f(x)dx = \int_0^{1/2} 2xdx = 1/4.$$

L–example 2.6 Let X be a random variable with probability function given by

$$f(x) = \begin{cases} c|x|, & |x| \leq 1 \\ 0, & \text{elsewhere.} \end{cases}$$

Find c.

Solution: The area under the curve $|x|, |x| \leq 1$ is $2 \times (1 \times 1/2) = 1$.



Therefore $c \cdot 1 = 1$ results in $c = 1$.

L–example 2.7

- "Time headway" in traffic flow is the elapsed time between the time that one car finishes passing a fixed point and the instant that the next car begins to pass that point.
- Let X = the time headway for two randomly chosen consecutive cars on a highway during a period of heavy flow.
- The following p.d.f. for X was suggested:

$$f(x) = \begin{cases} 0.15e^{-0.15(x-0.5)}, & \text{for } x \geq 0.5; \\ 0, & \text{otherwise.} \end{cases}$$

(a) Verify that $f(x)$ is a legitimate p.d.f. for the RV X.

(b) Compute $P(X \leq 5)$.

Solution:

(a) To check that $f(x)$ is a p.d.f., we need only to verify (1) $f(x) \geq 0$ for any $x \in \mathbb{R}$; (2) $\int_{-\infty}^{\infty} f(x)dx = 1$. (1) is clearly satisfied, we prove (2):

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_{0.5}^{\infty} 0.15e^{-0.15(x-0.5)}dx \\ &= 0.15e^{0.075} \int_{0.5}^{\infty} e^{-0.15x}dx \\ &= 0.15e^{0.075} \left(-\frac{1}{0.15}e^{-0.15x}\right)\Big|_{0.5}^{\infty} = 1. \end{aligned}$$

(b)

$$\begin{aligned} P(X \leq 5) &= \int_{-\infty}^{5} f(x)dx = \int_{0.5}^{5} 0.15e^{-0.15(x-0.5)}dx \\ &= 0.15e^{0.075} \left(-\frac{1}{0.15}e^{-0.15x}\right)\Big|_{0.5}^{5} \\ &= e^{0.075}\left(-e^{-0.75} + e^{-0.075}\right) = 0.4908. \end{aligned}$$

3 Cumulative Distribution Function

Definition 1

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

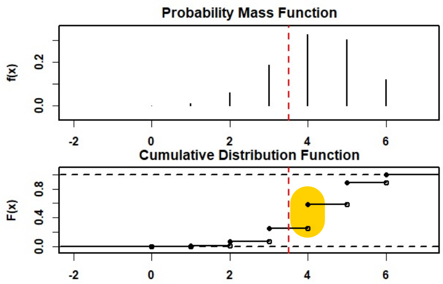
$$F(x) = P(X \leq x).$$

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

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



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

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

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

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

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

$\{X \le b\} = \{X < a\} \cup \{a \le X \le b\}$

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

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

EXAMPLE 2

- Revisit Examples 2 and 3. 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.:

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

EXAMPLE 3

Take the c.d.f. derived from Example 2:

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

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

<u>Solution</u>:

- 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

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

L–example 2.8

- Let $X =$ the number of days of sick leave taken by a randomly selected employee of a large company during a particular year.
- If the maximum number of allowable sick leave days per year is 14, possible values of X are $0, 1, 2, \ldots, 14$.
- Suppose $F(0) = 0.58$, $F(1) = 0.72$, $F(2) = 0.76$, $F(3) = 0.81$, $F(4) = 0.88$, and $F(5) = 0.94$.
- We have

$$\begin{aligned} P(2 \leq X \leq 5) &= F(5) - F(2-) \\ &= F(5) - F(1) = 0.94 - 0.72 = 0.22. \end{aligned}$$

- and

$$\begin{aligned} P(X=3) &= F(3) - F(3-) = F(3) - F(2) \\ &= 0.81 - 0.76 = 0.05. \end{aligned}$$

L–example 2.9 The p.f. for RV X is given by

$$f(x) = \begin{cases} (1-p)^{x-1} p, & \text{for } x = 1, 2, 3, \ldots; \\ 0, & \text{otherwise}, \end{cases}$$

where $p \in (0, 1)$ is a fixed value. Find the c.d.f. for X.
<u>Solution</u>:

- For any $x = 1, 2, 3, \ldots$, set $q = 1 - p$

$$\begin{aligned} F(x) &= P(X \leq x) = \sum_{t \leq x} f(t) = \sum_{t=1}^{x} (1-p)^{t-1} p \\ &= p\left(1 + q + q^2 + \ldots + q^{x-1}\right) \\ &= p \cdot \frac{1-q^x}{1-q} = 1 - (1-p)^x. \end{aligned}$$

- Question: What is the value of $F(x)$, when x is not a positive integer? For example, $x = 4.3$.

L–example 2.10 Suppose that the c.d.f. for RV X is given by

$$F(x) = \begin{cases} 1-(1-p)^{\lfloor x \rfloor}, & \text{for } x \geq 1; \\ 0, & \text{for } x < 1, \end{cases}$$

where $\lfloor x \rfloor$ denotes the integer part of x. For example, $\lfloor 3.6 \rfloor = 3$, $\lfloor 4 \rfloor = 4$, $\lfloor 4.7 \rfloor = 4$. Find its p.f. $f(x)$.
Solution:

- $F(x)$ changes values only for $x = 1, 2, 3, \ldots$; therefore it is a discrete distribution.
- $R_X = \{1, 2, 3, \ldots, \}$, i.e., the set of positive integers.
- for any $x \in R_X$,

$$\begin{aligned} f(x) &= F(x) - F(x-) = (1-(1-p)^x) - \left(1-(1-p)^{x-1}\right) \\ &= (1-p)^{x-1}(1-(1-p)) = (1-p)^{x-1}p, \end{aligned}$$

and $f(x) = 0$ otherwise.

L–example 2.11

- Many manufacturers have quality control programs that include inspection of incoming materials for defects.
- Suppose a computer manufacturer receives computer boards in lots of five.
- Two boards are selected from each lot for inspection.

(a) List all possible inspected boards for a lot.

(b) Suppose that boards 1 and 2 are the only defectives in a lot of five.

Define X = # of defective boards observed among an inspection.

Find the probability distribution of X.

(c) Let $F(x)$ be the c.d.f. of X. Derive $F(x)$.

Solution:

(a) $\#(S) = \binom{5}{2} = 10$. The possible selections are

$$\Big\{\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{2,3\},\{2,4\},\{2,5\},\{3,4\},\{3,5\},\{4,5\}\Big\}.$$

(b) X may take values of 0, 1, and 2.

$$\begin{aligned} f(0) &= P(X=0) = P(\{\{3,4\},\{3,5\},\{4,5\}\}) = 3/10, \\ f(2) &= P(X=2) = P(\{\{1,2\}\}) = 1/10, \\ f(1) &= P(X=1) = 1-[f(0)+f(2)] = 6/10, \end{aligned}$$

and $f(x) = 0$ elsewhere.

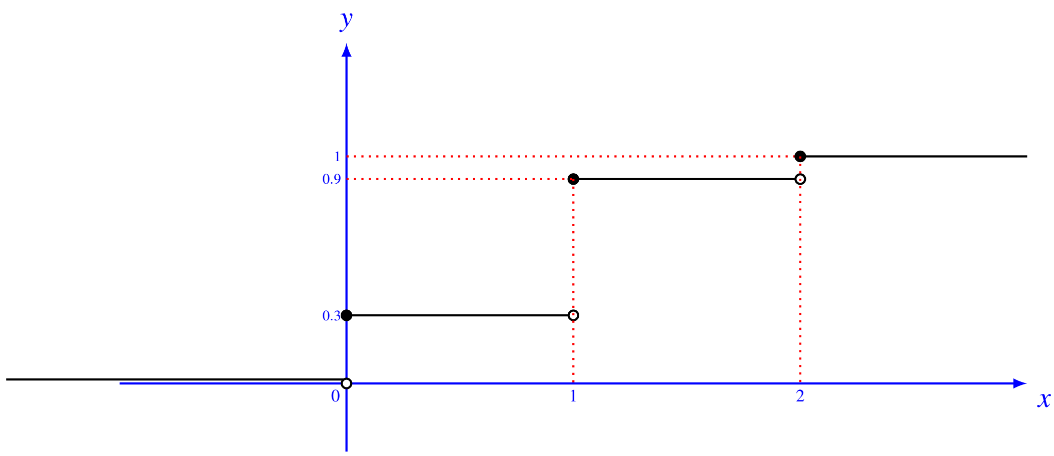
(c) It is sufficient to derive $F(0)$, $F(1)$, $F(2)$:

$$\begin{aligned} F(0) &= P(X \le 0) = f(0) = 0.3, \\ F(1) &= P(X \le 1) = f(0)+f(1) = 0.3+0.6 = 0.9 \\ F(2) &= P(X \le 2) = f(0)+f(1)+f(2) = 1. \end{aligned}$$

Therefore

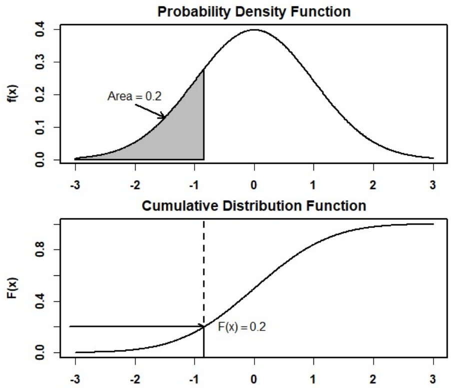
$$F(x) = \begin{cases} 0, & x < 0, \\ 0.3, & 0 \le x < 1, \\ 0.9, & 1 \le x < 2, \\ 1, & 2 \le x. \end{cases}$$

This c.d.f. can be drawn as a figure below:



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 4

- The p.d.f. of a RV X is given by
$$f(x) = \begin{cases} 2x & 0 \leq x < 1 \\ 0 & \text{elsewhere} \end{cases}$$
- The c.d.f. of X is
$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(t)dt \\ &= \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases} \end{aligned}$$

EXAMPLE 5

Take the c.d.f. derived from Example 4:
$$F(x) = \begin{cases} 0 & x < 0 \\ x^2 & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases}$$

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

L–example 2.12

- Let X be the vibratory stress (psi) on a wind turbine blade at a particular wind tunnel.
- The following p.d.f. for X is proposed:

$$f(x) = \begin{cases} \dfrac{x}{\theta^2} e^{-x^2/(2\theta^2)}, & \text{for } x > 0, \\ 0, & \text{otherwise,} \end{cases}$$

  where θ is a given constant.
- Verify that $f(x)$ is a legitimate p.d.f., and find its c.d.f. $F(x)$.

Solution:

- We first verify that $f(x)$ is a p.d.f.. It is obvious that $f(x) > 0$ for $x > 0$.

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)dx &= \int_0^{\infty} \frac{x}{\theta^2} e^{-x^2/(2\theta^2)} dx = -\int_0^{\infty} d\left(e^{-x^2/(2\theta^2)}\right) \\ &= -e^{-x^2/(2\theta^2)}\Big|_{x=0}^{\infty} \\ &= -0 - (-1) = 1. \end{aligned}$$

  This verifies that $f(x)$ is a valid p.d.f.
- For $x \le 0$, it is clearly $F(x) = 0$. For $x > 0$,

$$\begin{aligned} F(x) &= P(X \le x) = \int_{-\infty}^{x} f(t)dt = \int_0^x \frac{t}{\theta^2} e^{-t^2/(2\theta^2)} dt \\ &= -e^{-t^2/(2\theta^2)}\Big|_{t=0}^{x} \\ &= 1 - e^{-x^2/(2\theta^2)}. \end{aligned}$$

L–example 2.13 With the c.d.f. given in the last example:

$$F(x) = 1 - e^{-x^2/(2\theta^2)},$$

for $x \ge 0$ and $F(x) = 0$ otherwise. Derive its p.f.

- As $F(x)$ assumes different values in the interval $x \ge 0$, therefore we have continuous distribution. For any $x \ge 0$, we have

$$\begin{aligned} f(x) &= \frac{dF(x)}{dx} = \frac{d\left[1 - e^{-x^2/(2\theta^2)}\right]}{dx} \\ &= \frac{-d\left[e^{-x^2/(2\theta^2)}\right]}{dx} = \frac{x}{\theta^2} e^{-x^2/(2\theta^2)}, \end{aligned}$$

  and $f(x) = 0$ for $x < 0$ since $d(F(x))/dx = d(0)/dx = 0$. This complies with the p.d.f. given in the last example.

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 1 (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 2 (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 3
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 4
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

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

L–example 2.14

In a gambling game

- a man gains 5 if he gets all heads or all tails in tossing a fair coin 3 times;
- he pays 3 if either 1 or 2 heads show.

What is his expected gain?
Solution:

- Let X be the amount he can gain in the game.
- Then $X=5$ or -3 with the following probabilities:

$$\begin{aligned}P(X=5) &= P(\{HHH,TTT\})=1/8+1/8=1/4;\\ P(X=-3) &= 1-P(X=5)=3/4.\end{aligned}$$

- $E(X)=5\left(\frac{1}{4}\right)+(-3)\left(\frac{3}{4}\right)=-1.$
- This means he will lose 1 per toss, if he **continuously play the game for a long run**.

L–example 2.15

- Suppose "$X =$ the total number of hours (in units of 100 hours) that a family runs a vacuum cleaner over a period of one year".

- The probability function of X is given by

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

- Find the average number of hours per year that families run their vacuum cleaners.

Solution: The question is asking $100 \times E(X)$.

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x f(x) dx = \int_0^1 x \cdot x dx + \int_1^2 x(2-x) dx \\ &= \left(\frac{x^3}{3}\right)\Bigg|_0^1 + \left(x^2 - \frac{x^3}{3}\right)\Bigg|_1^2 \\ &= \left(\frac{1}{3} - 0\right) + \left[\left(4 - \frac{8}{3}\right) - \left(1 - \frac{1}{3}\right)\right] = 1. \end{aligned}$$

We conclude that on average, families run their vacuum cleaners 100 hours per year.

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,

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

L–example 2.16 Let X be a random variable, and let a and b be any real numbers. Show that

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

Solution:

- When X is a discrete random variable with p.f. $f(x)$,

$$\begin{aligned} E(aX+b) &= \sum_{x\in R_X}(ax+b)f(x) \\ &= \sum_{x\in R_X} axf(x) + \sum_{x\in R_X} bf(x) \\ &= a\left(\sum_{x\in R_X} xf(x)\right) + b\left(\sum_{x\in R_X} f(x)\right) = aE(X)+b. \end{aligned}$$

- When X is a continuous random variable with p.f. $f(x)$,

$$\begin{aligned} E(aX+b) &= \int_{-\infty}^{\infty}(ax+b)f(x)dx \\ &= \int_{-\infty}^{\infty}(ax)f(x)dx + \int_{-\infty}^{\infty} bf(x)dx \\ &= a\int_{-\infty}^{\infty} xf(x)dx + b\int_{-\infty}^{\infty} f(x)dx = aE(X)+b \end{aligned}$$

Note that based on properties (1) and (2), we have for constants $a_1, a_2, \ldots, a_k$ and RVs $X_1, X_2, \ldots, X_k$,

$$E(a_1X_1 + a_2X_2 + \ldots + a_kX_k) = a_1E(X_1) + a_2E(X_2) + \ldots + a_kE(X_k).$$

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 6
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 7
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

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

Compute $E(X)$, $V(X)$, and σ_X.
Solution:

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

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

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

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

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

L–example 2.17 Revisit Example 6. Let the p.f. of a RV X be given by

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

(a) Compute $V(X)$ with the alternative formula.

(b) Define $Y = X^2 + 2$. Compute $E(Y)$ and $V(Y)$.

Solution:

(a) We shall use the formula $V(X) = E(X^2) - [E(X)]^2$ to compute the variance. We can use $E(X) = 1$.

$$\begin{aligned} E(X^2) &= \sum_{x \in R_X} x^2 f(x) \\ &= (-1)^2\left(\frac{1}{8}\right) + 0^2\left(\frac{2}{8}\right) + 1^2\left(\frac{1}{8}\right) + 2^2\left(\frac{4}{8}\right) = 9/4. \\ V(X) &= E(X^2) - [E(X)]^2 = 9/4 - 1^2 = 5/4. \end{aligned}$$

(b) $E(Y) = E(X^2) + 2 = 9/4 + 2 = 17/4$. We use $V(Y) = E(Y^2) - [E(Y)]^2$ to compute the variance.

$$\begin{aligned} E(Y^2) &= E[(X^2+2)^2] = E(X^4 + 4X^2 + 4) \\ &= E(X^4) + 4(9/4) + 4 = E(X^4) + 13 \\ &= (-1)^4\left(\frac{1}{8}\right) + 0^4\left(\frac{2}{8}\right) + 1^4\left(\frac{1}{8}\right) + 2^4\left(\frac{4}{8}\right) + 13 \\ &= 85/4; \end{aligned}$$

Therefore

$$V(Y) = E(Y^2) - [E(Y)]^2 = 85/4 - (17/4)^2 = 51/16.$$

L–example 2.18 Show the property of variance:

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

Solution:

$$\begin{aligned} V(X) &= E[(X-\mu_X)^2] \\ &= E(X^2 - 2X\mu_X + \mu_X^2) \\ &= E(X^2) - E(2X\mu_X) + E(\mu_X^2) \\ &= E(X^2) - 2\mu_X E(X) + \mu_X^2 \\ &= E(X^2) - 2\mu_X^2 + \mu_X^2 = E(X^2) - \mu_X^2, \end{aligned}$$

since $\mu_X = E(X)$ is a constant.

L–example 2.19 Show the property of the variance: $V(aX+b) = a^2V(X)$, where a and b are constants.

Solution: Note that this property is equivalent to the following two properties

(a) $V(aX) = a^2V(X)$, and

(b) $V(X+b) = V(X)$.

Therefore, we only need to show (a) and (b). For (a)

$$\begin{aligned} V(aX) &= E[(aX)^2] - [E(aX)]^2 = E(a^2X^2) - [aE(X)]^2 \\ &= a^2E(X^2) - a^2[E(X)]^2 = a^2V(X). \end{aligned}$$

For (b),

$$\begin{aligned} V(X+b) &= E[(X+b)^2] - [E(X+b)]^2 \\ &= E(X^2+2Xb+b^2) - [E(X)+b]^2 \\ &= E(X^2) + 2bE(X) + b^2 - \{[E(X)]^2 + 2bE(X) + b^2\} \\ &= E(X^2) - [E(X)]^2 = V(X). \end{aligned}$$

L–example 2.20 Suppose that RV X has p.f.

$$f(x) = \begin{cases} \dfrac{x}{225}, & 0 < x < 15 \\ \dfrac{30-x}{225}, & 15 \le x \le 30 \\ 0, & \text{otherwise} \end{cases}$$

Compute $E(X)$ and $V(X)$.

Solution:

$$\begin{aligned}
E(X) &= \int_0^{15} x\left(\frac{x}{225}\right)dx + \int_{15}^{30} x\left(\frac{30-x}{225}\right)dx \\
&= \frac{1}{225}\left\{\left(\frac{x^3}{3}\right)\Big|_0^{15} + \left(15x^2 - \frac{x^3}{3}\right)\Big|_{15}^{30}\right\} \\
&= \frac{1}{225}\left\{\frac{15^3}{3} + \left(15(30)^2 - \frac{30^3}{3} - 15(15)^2 + \frac{15^3}{3}\right)\right\} = 15.
\end{aligned}$$

$$\begin{aligned}
E(X^2) &= \int_0^{15} x^2\left(\frac{x}{225}\right)dx + \int_{15}^{30} x^2\left(\frac{30-x}{225}\right)dx \\
&= \frac{1}{225}\left\{\left(\frac{x^4}{4}\right)\Big|_0^{15} + \left(10x^3 - \frac{x^4}{4}\right)\Big|_{15}^{30}\right\} = \frac{525}{2} = 262.5.
\end{aligned}$$

Therefore

$$V(X) = E(X^2) - [E(X)]^2 = 262.5 - 15^2 = 37.5.$$