Lesson 2.1: Random Variable and Its Probability Distribution

Learning Outcomes

At the end of the lesson, students must be able to

1. Explain the intuitive and formal definitions of a random variable,
2. Determine if a function defined on a sample space is a random variable, and
3. Construct the probability mass function and cumulative distribution function of a discrete random variable.

1 Introduction

Random variables are central to the study of probability and statistics. Intuitively, a random variable assigns a numerical value to each outcome in a sample space (Ω). For example, if we toss a coin twice, then the sample space is given by $\Omega = \{HH, HT, TH, TT\}$. Suppose we are interested in the number of heads and label this variable as X. Then, for each outcome in Ω, X will take the following values:

$$
\begin{aligned}
HH &\rightarrow 2 \\
HT &\rightarrow 1 \\
TH &\rightarrow 1 \\
TT &\rightarrow 0
\end{aligned}
$$

In short, $X = 0, 1, 2$. Because the values that X takes on depends on the outcomes of a random experiment, then the variable X is assigned a special name. It is called a **random variable**.

In the succeeding discussion we provide a mathematical (formal) definitions of a random variable. In these definitions, $\mathscr{F}$ is a sigma-algebra of subsets of Ω. In other words, $\mathscr{F}$ is a collection of events defined on the sample space Ω. Again, P is the probability function defined on these events. The triple (Ω, $\mathscr{F}$, P) is called a **probability space**.

Definition 1:
Consider the probability space $(\Omega, \mathscr{F}, P)$. A random variable is a real-valued function on Ω , that is $X : \Omega \rightarrow \mathbb{R}$ such that for any Borel set B of the real numbers, the set $\{\omega : X(\omega) \in B\}$ belongs to $\mathscr{F}$ for every $\omega \in \Omega$.

A simplified version of this definition is given next.

Definition 2:
Consider the probability space $(\Omega, \mathscr{F}, P)$. Suppose X is a function from Ω to $\mathbb{R}$. Then X is a random variable, if for every $r \in \mathbb{R}$ and every $\omega \in \Omega$, the set $\{\omega : X(\omega) \leq r\} \in \mathscr{F}$.

The two definitions are equivalent, but we shall use the latter in showing that a function $X : \Omega \to \mathbb{R}$ is a random variable.

<u>Example 1</u>:

Consider the experiment of tossing a single coin. We have $\Omega = \{head, tail\}$. Let the variable X denote the number of heads as follows,

$$X(\omega) = 1, \text{if } \omega = head$$
$$X(\omega) = 0, \text{if } \omega = tail$$

Is X a random variable?

<u>Solution</u>:

A trivial sigma-algebra is given by $\mathscr{F} = \{\Omega, \{head\}, \{tail\}, \phi\}$. Now,

- if $r < 0$, then the event $\{\omega : X(\omega) \leq r\} = \phi \in \mathscr{F}$
- if $0 \leq r < 1$, then the event $\{\omega : X(\omega) \leq r\} = \{tail\} \in \mathscr{F}$
- if $r \geq 1$, then the event $\{\omega : X(\omega) \leq r\} = \{head, tail\} = \Omega \in \mathscr{F}$

Observe that for each $r \in \mathbb{R}$ the event $\{\omega : X(\omega) \leq r\} \in \mathscr{F}$, therefore, X is a random variable.

<u>Example 2</u>:

Suppose we have the finite sample space $\Omega = \{a, b, c, d\}$ and the sigma algebra $\mathscr{F} = \{\phi, \Omega, \{a, b\}, \{c, d\}\}$.

(a) Is the function defined as $X(a) = X(b) = 0$ and $X(c) = X(d) = 2$ a random variable?

(b) Is the function defined as $Y(a) = 0, Y(b) = 2, Y(c) = 4, Y(d) = 5$ a random variable?

<u>Solution</u>: (Left as a classroom exercise!)

Distribution function of a random variable

With every random variable we will associate its probability distribution, or simply distribution. The distribution of the random variable X refers to the assignment of probabilities to all events defined in terms of this random variable, that is, events of the form $\{\omega : X(\omega) \leq r, \forall r \in \mathbb{R}\}$

Definition:
The cumulative distribution function (CDF), or simply distribution function, of a random variable X, denoted by $F_X(x)$ is defined as the function with domain the real line and range the interval [0,1] which satisfies

$$F_X(x) = P(X \leq x) = P\left(\{\omega : X(\omega) \leq x\}\right), \forall x \in \mathbb{R}$$

Example 3:

Suppose that in tossing a coin a person stands to win Php2.00 if he rolls heads, and to loss Php1.50 if he rolls tails. Let X represent the winnings of the person on a toss.

(a) Show that X is a random variable.

(b) Find the distribution function of X, assuming that the probability of heads is 0.6.

Solution:

(a) The sample space is $\Omega = \{Head, Tail\}$. Recall that the power set of Ω given by the collection $\mathscr{F} = \{\phi, \Omega, Head, Tail\}$ is a trivial sigma algebra. Now, from the given information, we have

$$X(\omega) = \begin{cases} -1.5, & \text{if } \omega = Tail \\ 2.0, & \text{if } \omega = Head \end{cases}$$

Since X takes on two values so we divide the real number line into 3 sub-intervals as follows: (-$\infty, -1.5), [-1.5, 2.0), [2.0, \infty)$.

Then evaluate if for each sub-interval the event $\{\omega : X(\omega) \leq r, r \in \mathbb{R}\}$ belongs to $\mathscr{F}$. The details are shown below.

(i) if $r \in (-\infty, -1.5), \{X(\omega) \leq r\} = \phi$

(ii) if $r \in [-1.5, 2.0), \{X(\omega) \leq r\} = \{Tail\}$

(iii) if $r \in [2.0, \infty), \{X(\omega) \leq r\} = \{Head, Tail\} = \Omega$

In each sub-interval the event $\{\omega : X(\omega) \leq r, r \in \mathbb{R}\}$ belongs to $\mathscr{F}$. Therefore, X is a random variable

(b) From (a) and by the definition of the distribution function, we have

(i) $F_X(x) = P((\{\omega : X(\omega) \leq x\}) = P(\phi) = 0$

(ii) $F_X(x) = P((\{\omega : X(\omega) \leq x\}) = P(\{Tail\}) = 0.4$

(iii) $F_X(x) = P((\{\omega : X(\omega) \leq x\}) = P(\Omega) = 1$

To summarize, the CDF of X is given by

$$F_X(x) = \begin{cases} 0, \text{ if } x < -1.5 \\ 0.4, \text{ if } -1.5 \leq x < 2.0 \\ 1, \text{ if } x \geq 2.0 \end{cases}$$

Theorem:
Every CDF of a random variable X has the following properties:

1. $0 \leq F_X(t) \leq 1, \forall t \in mathbbR$
2. $F_X(x)$ is non-decreasing. That is, if $a \leq b$, then $F_X(a) \leq F_X(b)$
3. $\lim_{t \to \infty} F_X(t) = 1$ and $\lim_{t \to -\infty} F_X(t) = 0$
4. $F_X(x)$ is continuous from the right. That is $F_X(t^+) = F_X(t^-)$

The proof of this theorem is left as your reading assignment. You can find this in a lot of books in mathematical statistics or even in the internet.

2 Probability distribution of a discrete random variable

Definition:
The support or domain of a random variable X is set of all possible values that it can assume. A random variable X whose support is countable (finitely or infinitely) is called a **discrete** random variable.

Example 4:

Below are a few examples of discrete random variables and their domain.

1. The sum of the number of dots on the upturned faces when a pair of dice is rolled. The domain is the set $\{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$.
2. The number of heads when three coins are tossed. The domain is the set $\{0, 1, 2, 3\}$
3. The daily number of new COVID-19 cases in Eastern Visayas. The domain is the set $\{0, 1, 2, 3, \cdots\}$

In the above examples, it is clear that a discrete random variable assumes values corresponding to the natural numbers or counting numbers, otherwise called the whole numbers.

For purposes of convention, we will use an uppercase letter, such as X or Y, to denote a random variable and a lowercase letter, such as x or y, to denote a particular value that a random variable may assume. For example, we can denote the number of heads when three coins are tossed as X and one possible value is x=2. Similarly, we can use Y to symbolize the number of new COVID-19 cases in Eastern Visayas and in a day there could be y = 38 new cases.

Definition:
Suppose that X is a discrete random variable. The function $p_X(x) = P(X = x)$ is called the probability mass function (pmf) for X.

Remarks:

1. The pmf $_X(x)$ consists of two parts:
   (a) the domain of X, and
   (b) a probability assignment $P(X = x)$, for all $x \in \mathbb{R}$.
2. For any discrete probability distribution, the following must be true:
   (a) $p_X(x) = P(X = x) \geq 0$
   (b) $\sum_{\forall x} p_X(x) = 1$
3. The pmf of a discrete random variable can be presented as a table, a graph, or a formula.
4. Given the pmf of a discrete random variable we can derive its CDF, and vice versa.

Example 5:

Consider the random experiment of tossing three coins. The sample space is given by $\Omega = \{HHH, HHT, HTH, THH, TTH, THT, HTT, TTT\}$, where H denotes a Head and T a Tail. If the coins are all fair, then we can assume that the outcomes in Ω are equally likely to occur. Let Y be the random variable denoting the number of heads obtained in this experiment. Then the domain of Y is $\{0, 1, 2, 3\}$. The probabilities assigned to each value of Y in the domain are:

$$P(Y = 0) = P(\{TTT\}) = \frac{1}{8}$$
$$P(Y = 1) = P(\{HTT, THT, TTH\}) = \frac{3}{8}$$
$$P(Y = 2) = P(\{HHT, HTH, THH\}) = \frac{3}{8}$$
$$P(Y = 3) = P(\{HHH\}) = \frac{1}{8}$$

We can then summarize and display the possible values of Y and the corresponding probabilities in a table as follows:

Y	0	1	2	3
$P(Y = y)$	1/8	3/8	3/8	1/8

Table 1: PMF of Y in Table Form

Notice that all the probabilities in the second row are all positive and their sum is 1. We call this table as the probability distribution table of Y. This is the probability mass function of Y displayed in table form.

We can also display the same information in the above table as a (vertical bar) graph. The heights of the bars correspond to the probabilities assigned to each value of Y. This is shown in Figure 1.

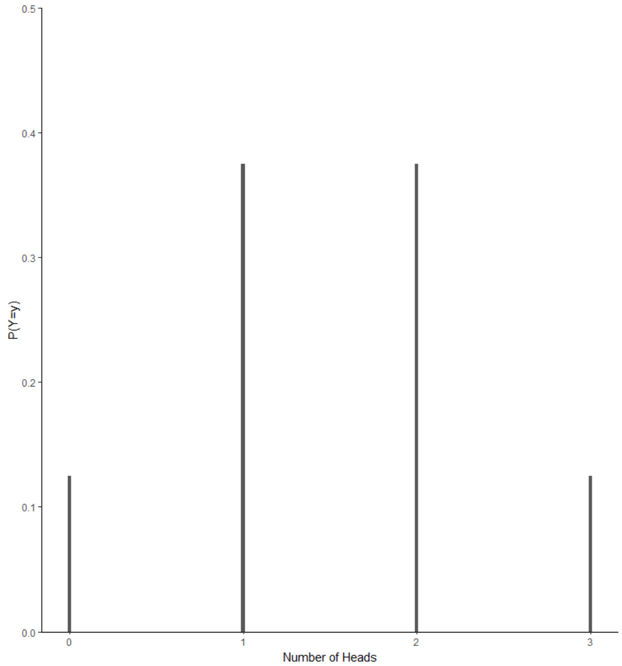

Figure 1: PMF of Y in Graphical Form

Finally, we can also express the probability distribution of Y as a formula and this is given below.

$$P(Y = y) = \binom{n}{y} p^y (1-p)^{n-y}$$

Generally, we call this distribution as the binomial distribution and we will discuss this in more detail in the later lessons.

From the probability mass function (pmf) of Y we can derive its cumulative distribution function (CDF). We do this using the definition of a CDF as follows:

1. If $y < 0$, say $y = -0.5$, then $F_Y(-0.5) = P(Y \leq -0.5) = 0$
2. If $0 \leq y < 1$, say $y = 0.7$, then $F_Y(0.7) = P(Y \leq 0.7) = P(Y = 0) = \frac{1}{8}$
3. If $1 \leq y < 2$, say $y = 1.8$, then $F_Y(1.8) = P(Y \leq 1.8) = P(Y = 1) + P(Y = 0) = \frac{3}{8} + \frac{1}{8} = \frac{1}{2}$
4. If $2 \leq y < 3$, say $y = 2.4$, then $F_Y(2.4) = P(Y \leq 2.4) = P(Y = 2) + P(Y = 1) + P(Y = 0) = \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = \frac{7}{8}$
5. If $y \geq 3$, say $y = 3.2$, then $F_Y(3.2) = P(Y \leq 3.2) = P(Y = 3) + P(Y = 2) + P(Y = 1) + P(Y = 0) = \frac{1}{8} + \frac{3}{8} + \frac{3}{8} + \frac{1}{8} = 1$

Therefore, the CDF is summarized as follows:

$$F_Y(y) = \begin{cases} 0, \text{ if} y < 0 \\ 1/8, \text{ if} 0 \leq y < 1 \\ 1/2, \text{ if} 1 \leq y < 2 \\ 7/8, \text{ if} 2 \leq y < 3 \\ 1, \text{ if} y \geq 3 \end{cases}$$

Example 6:

The CDF of a discrete random variable X is given by

$$F_X(x) = \begin{cases} 0,\ x < 0 \\ 1/5,\ 0 \le x < 1 \\ 4/5,\ 1 \le x < 2 \\ 1,\ x \ge 2 \end{cases}$$

Determine the pmf of X. Write this pmf in table form.

Solution:
It is obvious in the CDF that X assumes values of 0, 1, or 2 (refer to $\le$ or $\ge$ signs). Next, we find the probabilities assigned to each of these values.

$$P(X=0) = F_X(0 \le x < 1) - F_X(x < 0) = \frac{1}{5} - 0 = \frac{1}{5}$$
$$P(X=1) = F_X(1 \le x < 2) - F_X(0 \le x < 1) = \frac{4}{5} - \frac{1}{5} = \frac{3}{5}$$
$$P(X=2) = F_X(x \ge 2) - F_X = 1 - (1 \le x < 2) = 1 - \frac{4}{5} = \frac{1}{5}$$

We summarize all these findings in the following table.

X	0	1	2
$P(X=x)$	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

Example 7:

Determine the value of the constant k such that $f_X(x) = \frac{x}{k},\ x = 1, 2, 3, 4$ satisfies the conditions of a pmf.

Solution:

From (b) of Remark No. 2, $\sum_{\forall x} p_X(x) = 1$, That is,

$$\begin{aligned} \sum_{\forall x} p_X(x) = 1 &\implies \sum_{\forall x} \frac{x}{k} = 1 \\ &\implies \frac{1}{k} + \frac{2}{k} + \frac{3}{k} + \frac{4}{k} = 1 \\ &\implies k = 10 \end{aligned}$$

Therefore, the function $f_X(x) = \frac{x}{10},\ x = 1, 2, 3, 4$ is a valid pmf.

Example 8:

A supervisor in a manufacturing plant has three men and three women working for him. He wants to choose two workers for a special job. Not wishing to show any biases in his selection, he decides to select the two workers at random. Let X denote the number of women in his selection. Verify that the probability distribution for X is as follows:

X	0	1	2
$P(X = x)$	$\frac{1}{5}$	$\frac{3}{5}$	$\frac{1}{5}$

Solution: Left as a classroom exercise!

Example 9:

Consider the random experiment of tossing a pair of dice. Let Y be the random variable representing the sum of the number of dots on the sides facing up.

(a) Construct the probability mass function of Y.

(b) Based on (a), derive the CDF of Y.

(c) What is the probability that X is at most 9?

Solution: Left as a classroom exercise!