Stat 121 (Mathematical Statistics I)

Lesson 2.1. Random Variables and their Probability Distributions

Learning Outcomes

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

1. Explain the intuitive and formal definition 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.

What is a random variable?

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:

$$\begin{aligned} HH &\to X = 2 \\ HT &\to X = 1 \\ TH &\to X = 1 \\ TT &\to X = 0 \end{aligned}$$

In short, $X = 0, 1, 2$. Because the values that X takes on are random, the variable X has a special name. It is called a **random variable**. 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.

In the succeeding discussion we provide 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. P is the probability function defined on these events. The triple $(\Omega, \mathscr{F}, P)$ is called a *probability space*.

Definition

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

Definition

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

NOTE: 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 2.1.1</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,

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

So, the variable X assigns a real number with each outcome of the experiment. Suppose we consider a trivial sigma-algebra such as $\mathscr{F} = \{\phi, \Omega, \{head\}, \{tail\}\}$. Now,

1. If $r < 0$, then the event $\{\omega : X(\omega) \leq r\} = \phi$ since there is no outcome in Ω where X assigns a negative real number. Note that $\phi \in \mathscr{F}$.
2. If $0 \leq r < 1$, then $\{\omega : X(\omega) \leq r\} = \{tail\}$ since there is one outcome in Ω where X assigns a value of zero, which is $\omega = tail$. Note that $\{tail\} \in \mathscr{F}$.
3. If $r \geq 1$, then $\{\omega : X(\omega) \leq r\} = \{head, tail\} = \Omega$. Also, $\Omega \in \mathscr{F}$.

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

<u>Example 2.1.2</u>

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

1. Is the function X defined below a random variable?

$$X(\omega) = \begin{cases} 0, & \text{if } \omega = a,\ b \\ 2, & \text{if } \omega = c,\ d \end{cases}$$

2. Is the function Y defined below a random variable?

$$Y(\omega) = \begin{cases} 0, & \text{if } \omega = a \\ 2, & \text{if } \omega = b \\ 4, & \text{if } \omega = c \\ 5, & \text{if } \omega = d \end{cases}$$

Solution: Left as a classroom exercise!

Remark

Finding out that the event $\{\omega : X(\omega) \leq r\}$ does not belong to $\mathscr{F}$ in one sub-interval is already sufficient to conclude that a function is not a random variable.

Distribution function of a random variable

With every random variable we will associate its **probability 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 $\{\omega : X(\omega) \leq r\}$, for all $r \in \mathbb{R}$.

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

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

Example 2.1.3

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 represents the winnings of the person on a toss.

1. Show that X is a random variable.
2. Find the distribution function of X, assuming that the probability of heads is 0.6.

Solution

1. Left as a classroom exercise!
2. From the definition of the distribution function, we have

$$F_X(x) = P(\{\omega : X(\omega) \leq r\}) = P(\phi) = 0, \text{ if } x < -1.5$$
$$F_X(x) = P(\{\omega : X(\omega) \leq r\}) = P(\{tail\}) = 0.4, \text{ if } -1.5 \leq x < 2.0$$
$$F_X(x) = P(\{\omega : X(\omega) \leq r\}) = P(\Omega) = 1, \text{ if } x \geq 2.0$$

To summarize, the CDF of X is given by

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

Theorem

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

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

NOTE: 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.

Probability mass function (PMF) 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 2.1.4

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 total number of gadgets owned by students is the set $\{1, 2, 3, 4, \ldots\}$.

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.

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 $p_X(x)$ consists of two parts:

    a. the domain of X, and
    b. a probability assignment $P(X = x), \forall x \in \mathbb{R}$.

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.

<u>Example 2.1.5</u>

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 number of heads recorded in this experiment. Then the domain of Y is . The probabilities assigned to each value of Y in the domain are:

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

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

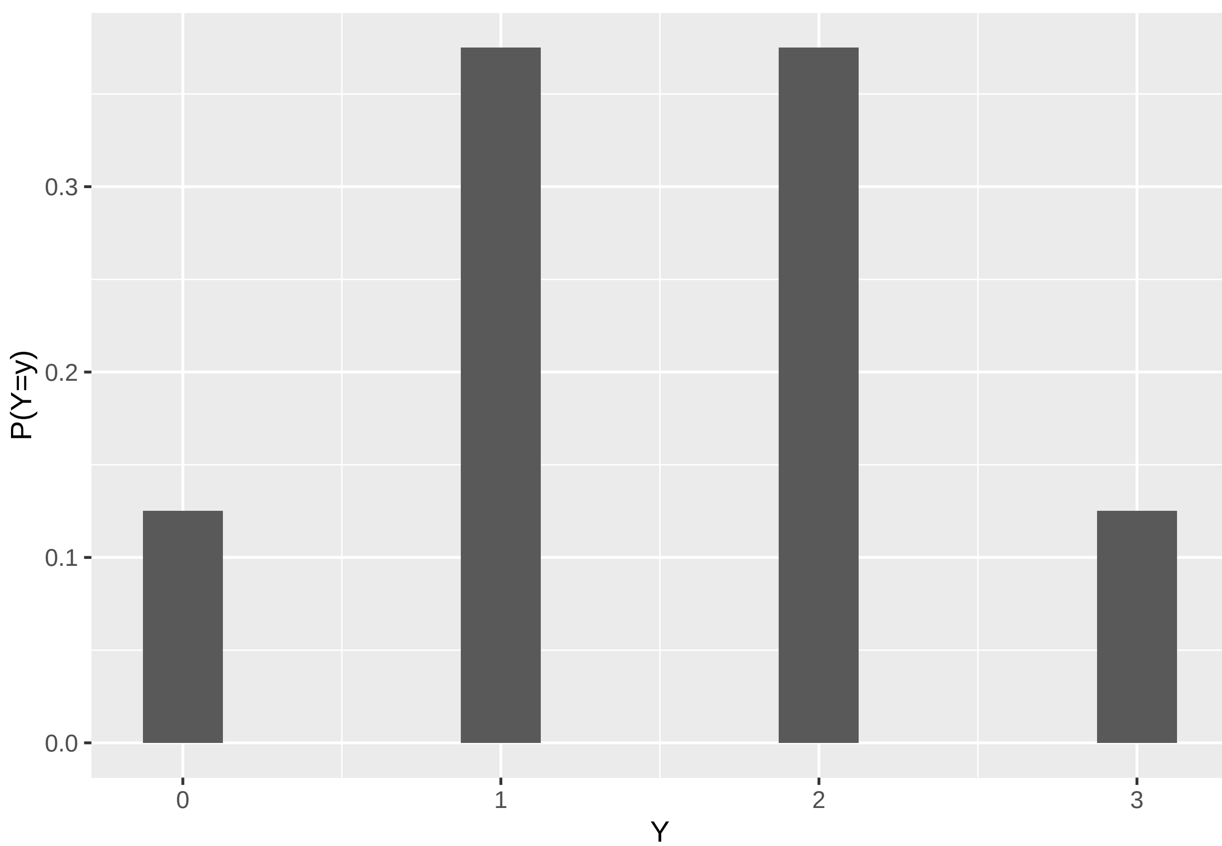
Y	0	1	2	3
P(Y=y)	1/8	3/8	3/8	1/8

Notice that all the probabilities (the entries in the 2nd row) are all positive and their sum is 1. This is the probability mass function of Y displayed in table form.

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

$$P(Y = y) = \binom{3}{y} 0.5^y (1 - 0.5)^{3-y}$$

Finally, the PMF can also be presented in graphical form.



Generally, the probability distribution that applies in this example is referred to as the binomial distribution and we will discuss this in more detail in the later lessons.

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

1. If $y < 0$, say $y = -0.5$, then $F_Y(-0.5) = P(Y \leq -0.5) = 0$, since there is no value of Y that is negative.

2. If $0 \leq y < 1$, say $y = 0.75$, then $F_Y(0.75) = P(Y \leq 0.75) = P(Y = 0) = 1/8$.

3. If $1 \leq y < 2$, say $y = 1.3$, then $F_Y(1.3) = P(Y \leq 1.3) = P(Y = 0) + P(Y = 1) = 1/8 + 3/8 = 1/2$.

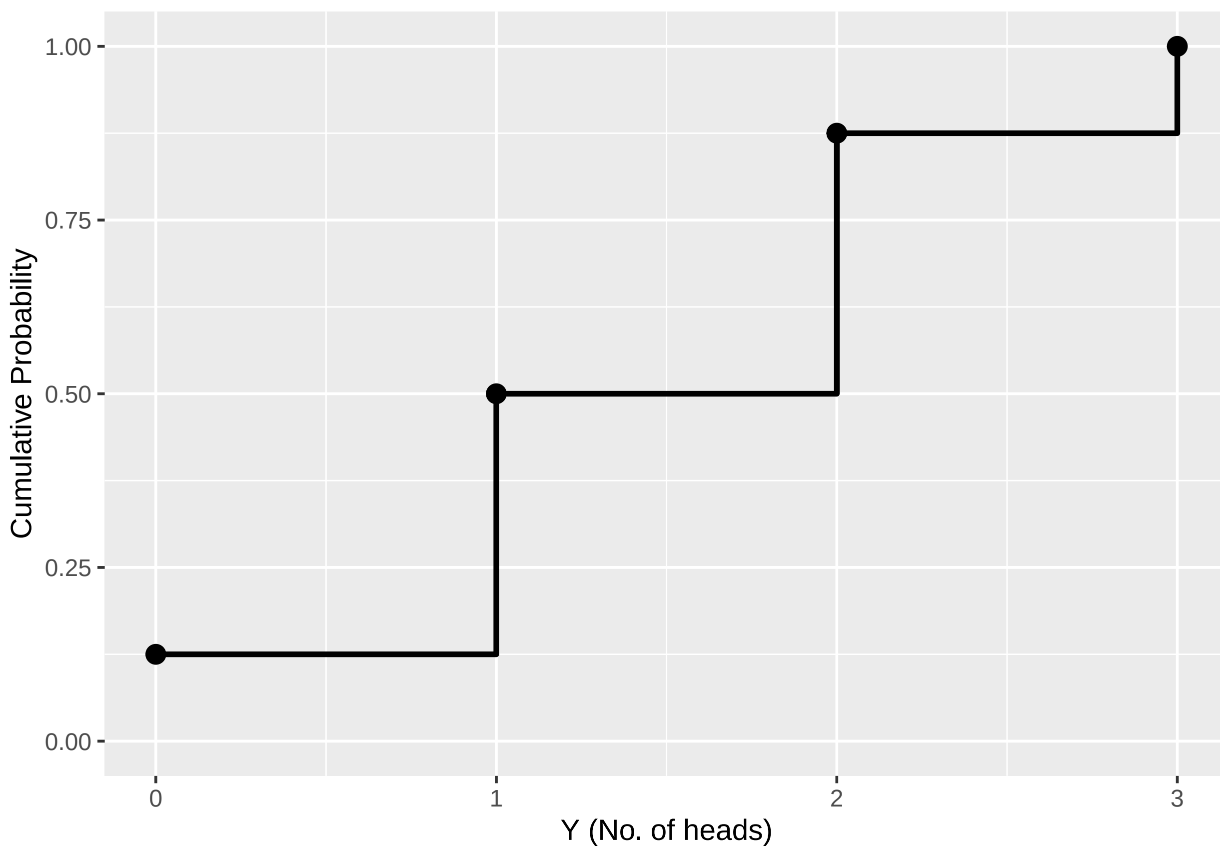
4. If $2 \leq y < 32$, say $y = 2.2$, then $F_Y(2.2) = P(Y \leq 2.2) = P(Y = 0) + P(Y = 1) + P(Y = 2) = 1/8 + 3/8 + 3/8 = 7/8$.

5. If $y \geq 3$, say $y = 3.6$, then $F_Y(3.6) = P(Y \leq 3.6) = P(Y = 0) + P(Y = 1) + P(Y = 2) + P(Y = 3) = 1/8 + 3/8 + 3/8 + 1/8 = 1$.

In summary, the CDF of Y is given by

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

The graph of this CDF is shown below.



Example 2.1.6

The CDF of a discrete random variable X is given by

$$F_X(x) = \begin{cases} 0, \text{ if } x < 0 \\ 1/5, \text{ if } 0 \leq x < 1 \\ 4/5, \text{ if } 1 \leq x < 2 \\ 1, \text{ if } x \geq 2 \end{cases}$$

Determine the PMF of X and write it in table form.

Solution

$$\begin{aligned} P(X = 0) &= F_X(0 \leq x < 1) - F_X(x < 0) = \frac{1}{5} - 0 = 1/5 \\ P(X = 1) &= F_X(1 \leq x < 2) - F_X(0 \leq x < 1) = \frac{4}{5} - \frac{1}{5} = \frac{3}{5} \\ P(X = 2) &= F_X(x \geq 2) - F_X(1 \leq x < 2) = 1 - \frac{4}{5} = \frac{1}{5} \end{aligned}$$

X	0	1	2
P(X = x)	1/5	3/5	1/5

Example 2.1.7

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

Solution

Recall that $\sum_{\forall x} p_X(x) = 1$. This means that

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

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

Example 2.1.8

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)	1/5	3/5	1/5

Solution

Since two workers are to be selected for a special job, the most number of women to be selected will be 2. In other words, the domain for X is the set $\{0, 1, 2\}$. Next we verify the probabilities.

$$P(X = 0) = \frac{\binom{3}{0}\binom{3}{2}}{\binom{6}{2}} = 1/5$$
$$P(X = 1) = \frac{\binom{3}{1}\binom{3}{1}}{\binom{6}{2}} = 3/5$$
$$P(X = 2) = \frac{\binom{3}{2}\binom{3}{0}}{\binom{6}{2}} = 1/5$$

Learning Tasks/Activities

1. Consider the experiment of tossing two coins. Let $Y(\omega) = 1$, if $\omega = HT$ or $\omega = TH$, and $Y(\omega) = 0$, otherwise. Is Y a random variable if $\mathscr{F} = \{\phi, \Omega, \{HH\}, \{HT, TH, TT\}\}$? Justify your answer.

2. 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 Y is at most 9?

3. The cumulative distribution function of a discrete random variable is given by

$$F_Y(y) = \begin{cases} 0, \text{ if } y < -2 \\ \frac{1}{12}, \text{ if } -2 \leq y < 2 \\ \frac{1}{3}, \text{ if } 2 \leq y < 5 \\ \frac{2}{3}, \text{ if } 5 \leq y < 6 \\ 1, \text{ if } y \geq 6 \end{cases}$$

Derive the probability mass function of Y.

4. Let a card be selected from an ordinary deck of playing cards. The outcome ω is one of 52 cards. Let $X(\omega) = 4$ if ω is an ace, $X(\omega) = 3$ if ω is a king, $X(\omega) = 2$ if ω is a queen, $X(\omega) = 1$ if ω is a jack, and $X(\omega) = 0$ for the rest of the cards.

   a. Construct the probability mass function of X.
   b. Determine the cumulative distribution function of X.
   c. What is the probability that X is less than 2?