

DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

TABLE OF CONTENTS

<i>Section Number and Heading</i>	<i>Page Number</i>
<i>Learning Objectives</i>	2
1. Random Experiments	2
2. Random Variables	4
3. Probability Distributions for Discrete Random Variables	7
4. Graphical Presentations of Probability Distributions	11
5. Parameters Of A Probability Distribution	12
6. The Cumulative Distribution Function	14
7. Deriving Probability Mass Function from Cumulative Distribution Function	16
<i>Practice Questions</i>	19

Content Developer

Chandra Goswami, Associate Professor, Department of Economics

Dyal Singh College, University of Delhi

Reference

Jay L. Devore: *Probability and Statistics for Engineering and the Sciences*, Cengage Learning, 8th edition [Chapter 3]

.

DISCRETE RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

Learning Objectives:

In this chapter you will learn what is a random variable and the two fundamentally different types of random variables. You will learn how to arrive at the probability distributions of discrete random variables and how to represent these graphically, as well as presentation by summary expressions. This provides the tool for evaluating the probability that the random variable takes on specific values or a range of values. You will also learn how the probability distribution can be used to specify a mathematical model for the population distribution. This will help you to identify the characteristics of the population. The chapter ends with practice questions so that you can test your understanding of the chapter contents.

Chapter Outline

1. Random experiments
2. Random variables
3. Probability distributions for discrete random variables
4. Graphical presentations of probability distributions
5. Parameters of a probability distribution
6. The cumulative distribution function for discrete random variables
7. Deriving probability mass function from cumulative distribution function

1. RANDOM EXPERIMENTS

A random or chance experiment is an experiment which yields different possible outcomes. These outcomes may be qualitative or quantitative. In case of qualitative outcomes, we observe a specific attribute of the variable. Quantitative outcomes result when we observe a number describing the attribute of the variable. Until the outcome is observed there is uncertainty about which particular outcome will be the result of the experiment. If the experiment is repeated under identical conditions different outcomes are likely to be observed at each trial.

Example 1.1

If a balanced coin is tossed there are two equally possible (qualitative) outcomes, a head (H) or a tail (T).

Example 1.2

It is known that wind speed and direction affects time taken by aircraft to reach their destination. The three possible outcomes for arrival time on any day are: before time, on time, or delayed.

Example 1.3

If an unbiased die is tossed it will result in one of six possible outcomes, depending on which face shows up: 1, 2, 3, 4, 5, or 6.

Example 1.4

If example 1.2 is restated to measure the extent of time delay in the aircraft reaching its destination, we can denote the possible outcomes as $x = 0$ for ontime arrival (ie, as per the scheduled time), $x < 0$ as measure of before time arrival (eg, - 5 minutes indicates arrival is 5 min ahead of scheduled time), and $x > 0$ for late arrival (eg, $x = 22$ represents arrival is 22 min after the scheduled time). We obtain an infinite number of possible outcomes since time is a continuous variable. Here extent of time delay (in minutes) is the variable where $x \in \mathbb{R}$.

Example 1.5

A bottling plant fills cold drinks in 200 ml bottles for its client. Although the machine is calibrated to dispense 200 ml per fill, it is noted that the fill amount varies from bottle to bottle by small amounts. If we denote $X =$ amount filled in a bottle (in ml), since volume is a continuous variable, the possible values of the variable are $x \in \mathbb{R}$

The outcomes in examples 1.1 and 1.2 are qualitative, and quantitative in examples 1.3, 1.4 and 1.5. There are a finite number of outcomes in examples 1.1, 1.2 and 1.3, whereas the number of outcomes is infinite in examples 1.4 and 1.5. In methods of statistical analysis we often need some numerical aspects of experimental outcomes. The mean, for instance, is a numerical function of the outcomes.

2. RANDOM VARIABLES

If the exhaustive set of all possible outcomes of a random experiment are known then probabilities of occurrence of the different outcomes can be assigned. The concept of a random variable allows us to obtain a numerical function of the experimental outcomes. Some of the most commonly used numerical functions of experimental outcomes are the mean, variance, and proportion.

Definition 1

For a given sample space S of some experiment, a *random variable* is any rule that associates a number with each outcome in S

A random variable (rv) is thus a function defined over the elements of S . The domain of the rv is the sample space and the range is a set of real numbers. A random variable is, therefore, a variable that takes on numerical values determined by the outcome of a random experiment. Thus, the value of the random variable will vary according to the observed outcome of a random experiment. In general, random variables are functions that associate numbers with some specific attribute of an experimental outcome. Random variables will be denoted by uppercase letters, such as X and Y , and their values by the corresponding lowercase letters, such as x and y .

Since the outcomes of a random experiment can be designated as a random variable, any numerical function of the outcomes is also a random variable. It is random since its value depends on which particular outcomes are observed. It is a variable since different numerical values are possible. We can, therefore, assign probabilities to its possible values. Therefore we can say that a random variable is a variable which can take one of the different possible values in the sample space with an assigned probability. If X denotes the rv and s the sample outcome, then $X(s) = q$ where q is a real number.

Example 2.1

If X is a rv with m possible values $x_1, x_2, x_3, \dots, x_m$ and Y is a rv with n possible values y_1, y_2, \dots, y_n then the linear function $X + Y$ is also a random variable since $x + y = x_i + y_j$ where $i = 1, 2, \dots, m$, and $j = 1, 2, \dots, n$.

Exercise 1

A balanced coin and a fair die are tossed simultaneously. List the different possible outcomes.

Solution:

Two possible outcomes of the coin are head (H) or tail (T). Six possible outcomes of the die are 1, 2, 3, 4, 5, and 6. Since the coin and die are tossed simultaneously the possible outcomes are as follows:

(H,1); (H,2); (H,3); (H,4); (H,5); (H,6); (T,1); (T,2); (T,3); (T,4); (T,5); (T,6)

Exercise 2

In Exercise 1, if H is denoted by 1 and tail by 0 so that $x = 0, 1$ and $y = 1, 2, 3, 4, 5, 6$ then list the different possible outcomes for the linear function $X+Y$

Solution:

$x + y = 1, 2, 3, 4, 5, 6, 2, 3, 4, 5, 6, 7$ according to the combinations listed in exercise 1

Exercise 3

Assigning appropriate probabilities to the values of the random variables X and Y in exercise 2, determine the probabilities of $x + y$.

Solution:

Since the coin is balanced, $P(x=0) = p(0) = \frac{1}{2}$ and $P(x=1) = p(1) = \frac{1}{2}$. Similarly, for the fair die, $p(1) = p(2) = p(3) = p(4) = p(5) = p(6) = \frac{1}{6}$.

Since X and Y are independent, there are 12 possible equally likely outcomes. Therefore, $p(1) = p(7) = \frac{1}{12}$ and $p(2) = p(3) = p(4) = p(5) = p(6) = \frac{2}{12}$

Definition 2

Any random variable whose only possible values are 0 and 1 is called a ***Bernoulli random variable***

If an unbiased coin is tossed repeatedly, on each toss there are only two possible outcomes so it is a Bernoulli rv. If an experiment can result in only two possible outcomes – success or failure – in each trial, we have a Bernoulli random variable.

There are fundamentally two different types of random variables: ***discrete random variables*** and ***continuous random variables***. The distinction between discrete and

continuous random variables lies in the number of possible values the rv can take. If the rv can have a finite number or a countably infinite number of possible values it is a discrete rv. If, on the other hand, the outcome can be any real number in a given interval, the number of possibilities is uncountably infinite, and the rv is said to be continuous.

Definition 3

A **discrete random variable** is a rv whose possible values either constitute a finite set or else can be listed in an infinite sequence which is “countably” infinite, where there is a first element, a second element, a third element and so on.

Examples 1.1, 1.2 and 1.3 have possible values which constitute a finite set. So is the case with exercise 1 and exercise.2. In all these cases the possible outcomes can be counted.

Example 2.2

A new company wishes to establish its brand image. For this purpose it runs a series of weekly newspaper advertisements until sales of its products reach the target level. Reaching the level of target sales is considered a success. Success may be achieved in 1 week or 2 weeks or 3 weeks and so on. If we denote success by S and failure by F then the sample space is $S = [S, FS, FFS, FFFS, \dots]$. We can define the random variable $X =$ number of weeks before the advertising campaign ends. Then, $X(S) = 1, X(FS) = 2, X(FFS) = 3, X(FFFS) = 4,$ and so on. Any positive integer is a possible value of X. Thus, the set of possible values of the rv X is countably infinite.

Variables which require counting, such as number of successes in an experiment, number of floors in a building, number of students in a class, shoe sizes, etc, are discrete.

Definition 4

A **random variable is continuous** if both the following conditions apply

1. Its set of possible values consists either of all numbers in an interval on the number line or all numbers in a disjoint union of such intervals.
2. No possible value of the random variable has a positive probability.

Condition 1 implies that there is no way to create a listing of all the infinite number of possible values of the variable. Condition 2 implies that intervals of values have positive probability. As the width of the interval diminishes, probability of the interval decreases. In the limit, probability of the interval is zero as the width of the interval reduces to zero.

Example 2.3

The university team is scheduled to visit any minute during a three hour long examination starting at 9am. We may want to find the probability that the team visits at a given time or we may be interested in the probability that the visit takes place during a given time interval. The sample space is from 0 to 180 minutes. The probability that the team visits during an interval of length c is $\frac{c}{180}$. This assignment of probabilities applies only to intervals on the measurement axis from 0 to 180. The probability decreases as the interval becomes shorter. For an interval of 5 seconds, the probability of a visit is $\frac{5}{10800} = 0.0004629$. As the length of the interval approaches zero, the probability that the team will visit also approaches zero. That is why we always assign zero probability for a single point on the number line. This does not mean that the team will not visit. The team will visit at some point in the interval from 0 to 180 even though each point has zero probability.

Variables such as time, height, distance, temperature, area, volume, weight, etc that require measurement are continuous. In practice, however, limitations of measurement instruments often do not allow measurement on a continuous scale. Yet it is useful to study models of continuous variables as they often reflect real world situations.

3 PROBABILITY DISTRIBUTIONS FOR DISCRETE RANDOM VARIABLES

A random experiment has different possible outcomes. It is certain that one of the possible outcomes will be observed as a result of the experiment. Let the experiment have only two possible outcomes which are mutually exclusive and exhaustive. Thus the probability that either one or the other outcome will occur is the sum of the probabilities. As there is no third alternative and the occurrence of one outcome precludes the occurrence of the other, the sum of the probabilities adds up to one. The probability

distribution of the rv X tells us how the total probability of 1 is distributed among the various possible values of the rv X . The probability assigned to any value x of the rv will be denoted by $p(x)$.

Definition 5

The probability distribution or ***probability mass function*** (pmf) of a discrete random variable is defined for every number x by $p(x) = P(X = x)$ for each x within the range of X .

Based on the postulates of probability, a function can serve as the pmf of X if and only if $p(x)$ satisfies the following two conditions

1. $0 \leq p(x) \leq 1$ for each value within its domain
2. $\sum_x p(x) = 1$ where the summation is over all values within its domain.

The first condition states that probability cannot be negative or exceed 1. The second condition follows from the fact that all possible values of X are mutually exclusive and collectively exhaustive so that the sum of the probabilities must equal 1. Thus, any function which satisfies both properties can serve as the pmf of a discrete random variable. Examples of pmf are Bernoulli Distribution, discrete Uniform Distribution, Binomial Distribution, Negative Binomial Distribution, Hypergeometric Distribution and Poisson Distribution.

Note that a function which satisfies the two conditions for one set of values of X may not do so for another set of values. In the latter case the function cannot serve as a pmf of X . To test whether a function is a pmf we need to check whether both conditions are satisfied for the given X values.

Exercise 4

A balanced coin is tossed three times. Let X denote the rv that is defined as the total number of heads. List the elements of the sample space and obtain the probability distribution of the total number of heads observed. Find a formula for the pmf of the total number of heads observed in three tosses of a fair coin.

Solution:

Denoting H = head and T = tail, elements of the sample space are

TTT, TTH, THT, HTT, THH, HTH, HHT, HHH.

Let the rv X = total number of heads observed in 3 tosses of a balanced coin. For a balanced coin a head and a tail are equally likely outcomes so that $P(H) = P(T) = \frac{1}{2}$. It can be assumed that the outcome of any toss is independent of the outcomes of the other two tosses of the coin. Then,

$$P(\text{TTT}) = P(X = 0) = p(0) = (1/2)(1/2)(1/2) = 1/8$$

$$P(\text{TTH or THT or HTT}) = P(X = 1) = p(1) = 3/8$$

$$P(\text{THH or HHT or HTH}) = P(X = 2) = p(2) = 3/8$$

$$P(\text{HHH}) = P(X = 3) = p(3) = 1/8$$

The probability distribution or pmf of X is given in the following table:

x	0	1	2	3
p(x)	1/8	3/8	3/8	1/8

The pmf can also be described as:

$$p(x) = \begin{cases} 1/8 & x = 0 \text{ or } 3 \\ 3/8 & x = 1 \text{ or } 2 \\ 0 & \text{otherwise} \end{cases}$$

Both conditions for a pmf are satisfied since $0 < p(x) < 1$ for $x = 0, 1, 2$ and 3 , and

$$\sum_x p(x) = 1$$

Based on the probabilities we observe that numerators of the four fractions $1/8, 3/8, 3/8$ and $1/8$ are the binomial coefficients $\binom{3}{0}, \binom{3}{1}, \binom{3}{2}, \binom{3}{3}$. The formula for the pmf can,

therefore, be written as $\frac{\binom{3}{x}}{8}$ for $x = 0, 1, 2$ and 3 .

Exercise 5

A computer shop sells desktops, laptops, notebooks and tablets. A prospective buyer enters the shop. The random variable can take five possible values. $X = 0$ if no purchase is made, $X = 1$ if a tablet is purchased, $X = 2$ if a notebook is purchased, $X = 3$ if a laptop is bought, and $X = 4$ if a desktop is bought. If 40% of buyers purchase a tablet, 35%

buyers opt for a notebook, 20% a laptop and 5% a desktop, what is the probability distribution of X?

Solution:

The pmf is as follows

x	0	1	2	3	4
p(x)	0	0.4	0.35	0.20	0.05

Exercise 6

A balanced coin is tossed four times. Use the formula derived in exercise 4 to obtain the pmf of X = total number of heads in four tosses of the coin.

Solution:

Total number of possible outcomes is $2^4 = 16$ as the result of each toss is independent of

the remaining three tosses. Using the formula $p(x) = \frac{\binom{4}{x}}{16}$ the pmf is as follows:

x	0	1	2	3	4
p(x)	1/16	4/16	6/16	4/16	1/16

Exercise 7

Check whether the function given by $f(x) = \frac{x+4}{30}$ for $x = 0, 1, 2, 3, 4$ can serve as the probability distribution of a discrete random variable.

Solution:

For given values of x the value of the function is as follows:

$$f(0) = 4/30, f(1) = 5/30, f(2) = 6/30, f(3) = 7/30, f(4) = 8/30$$

Each of the above values are positive fractions less than 1. Hence the first condition for a pmf is satisfied. The sum of all the values of f(x), $\sum f(x) = (4 + 5 + 6 + 7 + 8)/30 = 1$ so that the second condition is also satisfied. Since both the required conditions for a pmf are satisfied, therefore, the given function can serve as a pmf for a rv having the values 0, 1, 2, 3, and 4.

4 GRAPHICAL PRESENTATIONS OF PROBABILITY DISTRIBUTIONS

The pmf is positive for the countable number of values of the rv and zero for all other values. We have $\sum_x p(x) = 1$. Thus, the pmf describes how the total probability mass of 1 is distributed at various points on the number line. The pmf can be presented graphically in *probability histograms*.

For a probability histogram, above each x with $P(x) > 0$ construct a rectangle centered at x . The height of each rectangle is proportional to $P(x)$. The area of the rectangle equals $p(x_i)$ for $X = x_i$. If the base of each rectangle is of unit width then the height will be equal to $p(x_i)$ for $X = x_i$.

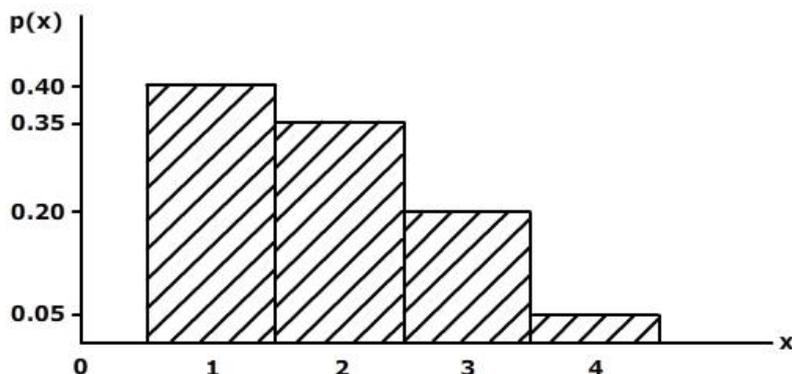
Example 4.1

The pmf of exercise 5 is

x	0	1	2	3	4
p(x)	0	0.40	0.35	0.20	0.05

For all $x > 4$, $p(x) = 0$. The probability histogram is drawn by representing 1 with the interval 0.5 to 1.5, 2 with the interval 1.5 to 2.5, 3 with the interval 2.5 to 3.5, and so on.

Figure 1 Probability histogram



The line graph and bar chart are also referred to as histograms. The line graph is drawn by drawing lines of height $p(x)$ for corresponding x values. The bar chart is drawn with each rectangle centered at the x value with a height equal to the probability of the corresponding value of the rv. The line graph and bar chart for the pmf of ex. 5 are illustrated in Fig 2 and Fig 3 respectively.

Figure 2 Line graph

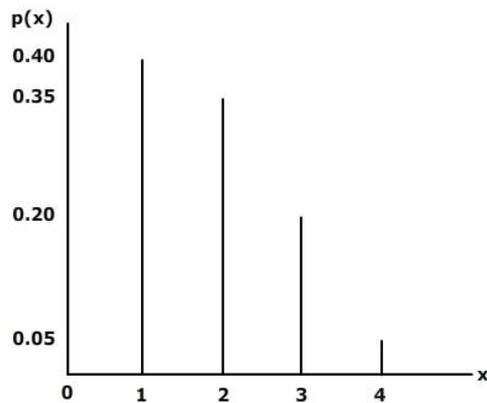
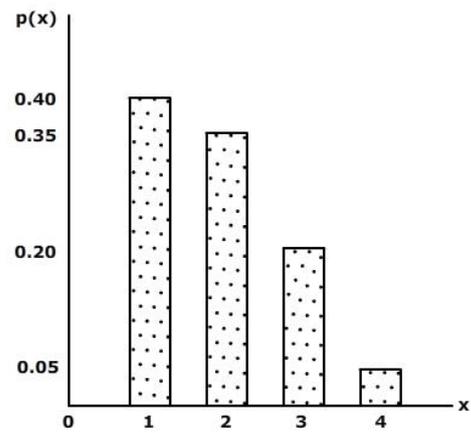


Figure 3 Bar chart



5 PARAMETERS OF A PROBABILITY DISTRIBUTION

We can use the pmf to specify a mathematical model for a discrete population distribution. If the population does not exist we can think of it as a mathematical model for a conceptual population. Let the population consist of X values $x_1, x_2, x_3, \dots, x_n$ with corresponding probabilities $p(x_i)$. From the relative frequency approach to probability, we know that limit of the relative frequency is the probability of occurrence of the X value in the population. As the population size tends to become infinitely large the relative frequency approaches the probability, ie,

$$\lim_{n \rightarrow \infty} \frac{f_i}{n} = p(x_i), \text{ where } f_i \text{ is the frequency of } x_i \text{ and } \sum f_i = n.$$

When all possible values x_i of the rv X are considered then $\sum p(x_i) = 1$ and we have the probability distribution for the discrete population. The pmf thus provides a model for the distribution of population values. Once we have such a population model we can use it to calculate the values of population characteristics, like the mean μ , variance σ^2 , etc, and make inferences about such characteristics.

Definition 6

Suppose $p(x)$ depends on a quantity that can be assigned any one of a number of possible values, with each different value determining a different probability distribution. Such a quantity is called a **parameter** of the distribution.

The collection of all probability distributions for different values of the parameter is called a **family** of probability distributions.

Example 5.1

We consider a random experiment that can give rise to just two possible mutually exclusive and exhaustive outcomes 0 and 1. Then $p(0) + p(1) = 1$. Such a rv is called a **Bernoulli** random variable. If we select α such that $0 < \alpha < 1$, the pmf of the Bernoulli rv can then be expressed as

$$p(x) = \begin{cases} 1 - \alpha & x = 0 \\ \alpha & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

For each of the possible values of α in the interval between 0 and 1, we obtain a different probability distribution. We thus obtain a family of Bernoulli distributions with each pmf determined by a particular value of α . Since the pmf depends on the particular value of α we often write the pmf of the Bernoulli distribution as $p(x; \alpha)$ rather than just $p(x)$. The quantity α in the Bernoulli pmf is a parameter. The value of the parameter α distinguishes one Bernoulli distribution from another. If α can take *any* value in the interval 0 to 1, we obtain an infinite number of Bernoulli distributions, each for a different value of α .

The value of the parameter may be unknown. If the population size is very large it may not be possible to examine all the population values to ascertain the value of α . We can then use sample data to infer about the parameter value α , where the sample is a representative subset of the population.

Example 5.2

If the discrete rv X can take any value $x_1, x_2, x_3, \dots, x_n$ with equal probability we have a discrete **Uniform Distribution**. We can denote the minimum value $x_1 = \alpha$, and the maximum value $x_n = \beta$. Then the pmf of the Uniform Distribution can be expressed as

$$p(x) = \begin{cases} \frac{1}{n} & \alpha \leq x \leq \beta \\ 0 & \text{otherwise} \end{cases}$$

We obtain a family of uniform distributions with the pmf of each distribution determined by a particular set of values for α and β . The pmf can be denoted by $p(x; \alpha, \beta)$, where α and β are the parameters of the distribution. For different combinations of values of α and β we obtain different uniform distributions.

6 THE CUMULATIVE DISTRIBUTION FUNCTION

It is often the case that we need to know the probability that the observed value of the rv is less than or equal to a specific value $x = a$. For this we require the cumulative distribution function (cdf) or, simply, the distribution function of the discrete rv X . It is also called the cumulative mass function. The cdf is denoted by $F(x)$.

Definition 7

The ***cumulative distribution function*** $F(x)$ of a discrete random variable X with pmf $p(x)$ is defined for every number x by $F(x) = P(X \leq x) = \sum_{y:y \leq x} p(y)$

Thus, cdf is obtained by summing the pmf $p(x)$ over all possible values of $X = y$ satisfying $y \leq x$. We use $F(x)$ to calculate the probability that the observed value of X does not exceed x . It follows that $P(X < x) < P(X \leq x)$ since the value x is included in $P(X \leq x)$ and not in $P(X < x)$. Only if $P(X = x) = 0$ then $P(X < x) = P(X \leq x)$. In all other cases where $P(X = x) > 0$ the inequality holds, ie, $P(X < x) < P(X \leq x)$.

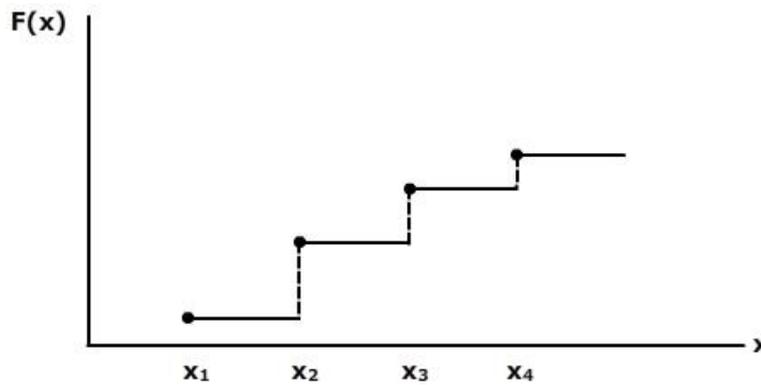
The cumulative distribution function has the following properties:

1. $0 \leq F(x_i)$ for every value $X = x_i$
2. If $a < b$ then $F(a) \leq F(b)$ where a and b are two possible values of the rv X .

The first property states that $F(x)$ is non-negative. $F(x) = 0$ for any value of X that is less than the smallest permissible X value of the pmf since $p(x) = 0$ for all such values. It follows that when all possible values of X have been considered $F(x) = 1$. For higher values of X we again have $p(x) = 0$ so that $F(x)$ remains unchanged at 1. The second property implies that if $p(b) = 0$ then $F(a) = F(b)$. Otherwise $F(a) < F(b)$ when $a < b$.

The graph of $F(x)$ is a step function. If X is a discrete rv whose set of possible values are x_1, x_2, \dots , where $x_1 < x_2 < x_3 < \dots$, the value of $F(x)$ is constant in the interval between two successive values x_{i-1} and x_i , and then increases by $p(x_i)$ at x_i . $F(x)$ again remains flat between x_i and x_{i+1} when it jumps up (takes a step) by $p(x_{i+1})$ at x_{i+1} . This is illustrated in Figure 4.

Figure 4 Graph of the cdf



Since $F(x_{i-1}) < F(x_i)$ and $F(x_i) < F(x_{i+1})$, at all points of discontinuity the cdf takes on the greater of the two values. This is indicated by heavy dots in Figure 4. It can be seen that as x increases, the cdf will change values only at those points that can be taken by the rv with positive probability.

Example 6.1

Using the pmf in exercise 5,

x	0	1	2	3	4
p(x)	0	0.40	0.35	0.20	0.05

$$F(0) = P(X = 0) = 0$$

$$F(1) = P(X = 0 \text{ or } 1) = 0 + 0.4 = 0.4$$

$$F(2) = P(X = 0 \text{ or } 1 \text{ or } 2) = 0.4 + 0.35 = 0.75$$

$$F(3) = P(X = 0 \text{ or } 1 \text{ or } 2 \text{ or } 3) = 0.75 + 0.20 = 0.95$$

$$F(4) = P(X= 0 \text{ or } 1 \text{ or } 2 \text{ or } 3 \text{ or } 4) = 0.95 + 0.05 = 1$$

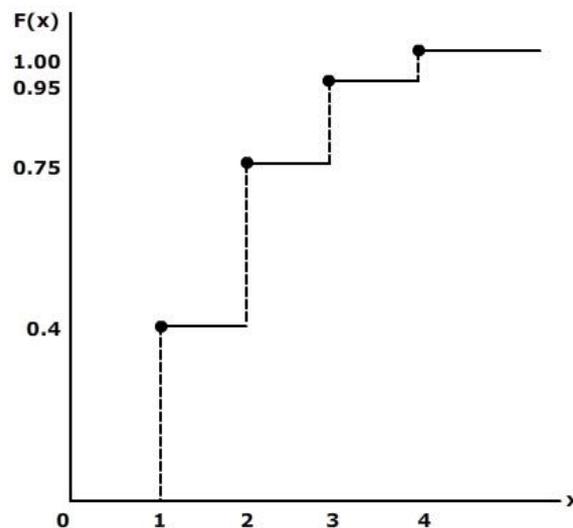
Since $P(x) = 0$ for all $x > 4$, $F(x > 4) = P(X \geq 4) = 1$

Hence, the cumulative distribution function of the rv X is as follows:

$$F(x) = \begin{cases} 0 & x=0 \text{ or } x < 1 \\ 0.40 & 1 \leq x < 2 \\ 0.75 & 2 \leq x < 3 \\ 0.95 & 3 \leq x < 4 \\ 1 & 4 \leq x \end{cases}$$

The graph for $F(x)$ is then as shown in Fig 5

Figure 5 Graph of the cdf for example 6.1



7 DERIVING PROBABILITY MASS FUNCTION FROM CUMULATIVE DISTRIBUTION FUNCTION

Just as the cdf can be derived from the pmf, if the cdf is known then the pmf can be derived from it. Also the probability that X falls in a specified interval can be obtained from the cdf. Suppose the range of a rv X consists of the values $x_1, x_2, x_3, \dots, x_n$ where $x_1 < x_2 < x_3, \dots, < x_n$, then

$$P(x_1) = F(x_1)$$

$$P(x_2) = F(x_2) - F(x_1)$$

$$P(x_3) = F(x_3) - F(x_2)$$

and so on. In general, $P(x_i) = F(x_i) - F(x_{i-1})$ for all $i = 1, 2, 3, \dots, n$. In this way we can derive the pmf from the cdf.

Example 7.1

Given the cdf obtained in example 6.1

$$F(x) = \begin{cases} 0 & x=0 \text{ or } x < 1 \\ 0.40 & 1 \leq x < 2 \\ 0.75 & 2 \leq x < 3 \\ 0.95 & 3 \leq x < 4 \\ 1 & 4 \leq x \end{cases}$$

we get

$$p(0) = 0$$

$$p(1) = 0.4 - 0 = 0.4$$

$$p(2) = 0.75 - 0.4 = 0.35$$

$$p(3) = 0.95 - 0.75 = 0.20$$

$$p(4) = 1 - 0.95 = 0.05$$

To obtain the probability that value of X falls in the interval $[a, b]$ such that $a \leq b$, where both a and b are included in the interval, we have to compute $P(a \leq X \leq b) = F(b) - F(a^-)$ where a^- denotes the largest possible X value that is strictly less than a . If the only possible values of X are integers so that a and b are both integers, then

$$P(a \leq X \leq b) = P(X = a \text{ or } a + 1 \text{ or } a + 2 \text{ or } \dots \text{ or } b) = F(b) - F(a - 1)$$

This principle can be used to find the probability that X takes the value a . By setting $b = a$ we obtain $P(X = a) = p(a) = F(a) - F(a - 1)$.

This method is used to derive the pmf from the cdf.

We can similarly compute $P(a < X < b) = F(b-1) - F(a)$ where a and b are not included in the interval.

Note that $F(b) - F(a)$ gives us $P(a < X \leq b)$ where b is included in the interval but a is not included.

Example 7.2

Given the cdf obtained in example 6.1,

$$P(1 \leq X \leq 3) = F(3) - F(0) = 0.95 - 0 = 0.95$$

$$P(1 < X \leq 3) = F(3) - F(1) = 0.95 - 0.4 = 0.55$$

$$P(X \geq 3) = 1 - P(X < 3) = 1 - F(2) = 1 - 0.75 = 0.25$$

Exercise 8

A study of number of delayed flights in an hour (X) at an airport due to fog in winter revealed the following probability distribution of the rv X .

x	0	1	2	3	4	5	6
$p(x)$	0.10	0.15	0.25	0.20	0.16	0.11	0.03

- (a) Derive the cdf
- (b) What is the probability at least three flights are delayed in an hour?
- (c) What is $P(1 < X < 5)$?

Solution

- (a) The cdf derived as follows:

$$F(0) = p(0) = 0.10$$

$$F(1) = F(0) + p(1) = 0.10 + 0.15 = 0.25$$

$$F(2) = F(1) + p(2) = 0.25 + 0.25 = 0.50$$

$$F(3) = F(2) + p(3) = 0.50 + 0.20 = 0.70$$

$$F(4) = F(3) + p(4) = 0.70 + 0.16 = 0.86$$

$$F(5) = F(4) + p(5) = 0.86 + 0.11 = 0.97$$

$$F(6) = F(5) + p(6) = 0.97 + 0.03 = 1.00$$

Since $X \geq 0$, the cdf can be presented as follows:

$$F(x) = \begin{cases} 0.10 & x=0 \text{ or } x < 1 \\ 0.25 & 1 \leq x < 2 \\ 0.50 & 2 \leq x < 3 \\ 0.70 & 3 \leq x < 4 \\ 0.86 & 4 \leq x < 5 \\ 0.97 & 5 \leq x < 6 \\ 1 & 6 \leq x \end{cases}$$

- (b) $P(X \geq 3) = 1 - F(2) = 1 - 0.50 = 0.50$
- (c) $P(1 < X < 5) = P(X = 2 \text{ or } 3 \text{ or } 4) = F(4) - F(1) = 0.86 - 0.25 = 0.61$

PRACTICE QUESTIONS

1. Suppose one die has spots 1, 2, 2, 3, 3, 4 and a second die has spots 1, 3, 4, 5, 6, 8. If both dice are rolled, list the sample space (all possible outcomes). Let the rv X = total number of spots showing. What is the pmf of X ? Show that this pmf is the same as that for two normal dice, each having 1, 2, 3, 4, 5, 6 spots.
2. At the points $x = 0, 1, 2, 3, 4, 5, 6$ the cdf for the discrete rv X is $F(x) = x(x+1)/42$. Find the pmf for X .
3. Urn 1 and urn 2 each have two red balls and two white balls. Two balls are drawn simultaneously from each urn. Let
 X_1 = number of red balls in the sample from first urn, and
 X_2 = number of red balls in the sample from the second urn.
Find the pmf of $X_1 + X_2$
4. An urn contains four balls numbered 1, 2, 3, and 4. If two balls are drawn from the urn at random and Z is the sum of the numbers on the two balls, find
 - (a) the probability distribution of Z and draw the histogram
 - (b) the cdf of Z and draw its graph
5. A coin is biased so that heads is twice as likely as tails. For three independent tosses of the coin, find
 - (a) the probability distribution of X , the total number of heads
 - (b) the probability of getting at most two heads, using the cdf of X
 - (c) $P(1 < X \leq 3)$ and $P(X > 2)$, using the cdf
6. The amount of coffee (in grams) in a 230-gm jar filled by a certain machine is a random variable whose probability density is given by

$$f(x) = \begin{cases} 0 & x \leq 227.5 \\ \frac{1}{5} & 227.5 < x < 232.5 \\ 0 & x \geq 232.5 \end{cases}$$

Find the probabilities that a 230-gram jar filled by this machine will contain

- (a) at most 228.65 gm of coffee
- (b) anywhere from 229.34 to 231.66 gm of coffee
- (c) at least 229.85 gm of coffee

7. A library subscribes to two different weekly news magazines, each of which is supposed to arrive in Wednesday's mail. In actuality, each one may arrive on Wednesday, Thursday, Friday, or Saturday. Suppose the two arrive independently of one another, and for each one $P(\text{Wed}) = 0.3$, $P(\text{Thu}) = 0.4$, $P(\text{Fri}) = 0.2$, and $P(\text{Sat}) = 0.1$. Let Y = the number of days beyond Wednesday that it takes for both magazines to arrive (so possible Y values are 0, 1, 2, or 3). List all the possible outcomes and compute the pmf of Y .

8. Given the following cdf, derive the pmf of Y and draw the

- (a) histogram of the pmf
- (b) graph of the cdf

$$F(y) = \begin{cases} 0 & y < 1 \\ 0.05 & 1 \leq y < 2 \\ 0.15 & 2 \leq y < 4 \\ 0.50 & 4 \leq y < 8 \\ 0.90 & 8 \leq y < 16 \\ 1 & 16 \leq y \end{cases}$$

9. A contractor is required by the city planning department to submit one, two, three, four, or five forms (depending on the nature of the project) in applying for a building permit. Let Y = the number of forms required of the next applicant. The probability that y forms are required is known to be proportional to y , ie, $p(y) = ky$ for $y = 1, 2, 3, 4, 5$

- (a) What is the value of k ?
- (b) What is the probability that at most three forms are required?
- (c) What is the probability that between two and four forms (inclusive) are required?

(d) Could $p(y) = \frac{y^2}{50}$ for $y = 1, \dots, 5$ be the pmf of Y ?

10. An insurance company offers its policyholders a number of different premium payment options. For a randomly selected policyholder, let X = the number of months between successive payments. The probability mass function of X is as follows:

x	1	3	4	6	12
p(x)	0.30	0.10	0.05	0.15	0.40

- (i) Derive the cumulative distribution function (cdf) of X and draw the graph of this cdf
- (ii) Using the cdf, compute $P(3 \leq X < 6)$, $P(3 \leq X \leq 6)$, and $P(4 \leq X)$.

