UNIT 10 DISCRETE PROBABILITY DISTRIBUTIONS

Objectives

After reading this unit, you should be able to :

- understand the concepts of random variable and probability distribution
- appreciate the usefulness of probability distribution in decision-making
- identify situations where discrete probability distributions can be applied
- find or assess discrete probability distributions for different uncertain situations
- appreciate the application of summary measures of a discrete probability distribution.

Structure

10.1 INTRODUCTION

In our study of Probability Theory, we have so far been interested in specific outcomes of an experiment and the chances of occurrence of these outcomes. In the last unit, we have explored different ways of computing the probability of an outcome. For example, we know how to calculate the probability of getting all heads in a toss of three coins. We recognise that this information on probability is helpful in our decisions. In this case, a mere 0.125 chance of all heads may dissuade you from betting on the event of "all heads". It is easy to see that it would have been further helpful, if all the possible outcomes of the experiment together with their chances of occurrence were made available. Thus, given your interest in betting on head's, you find that a toss of three coins may result in zero, one, two or three heads with the respective probabilities of $\frac{1}{8}$, $\frac{3}{8}$, $\frac{3}{8}$, and $\frac{1}{8}$. The wealth of information, presented in this way, helps you in drawing many different inferences. Looking at this information, you may be more ready to bet on the event that either one or two heads occur in a toss of three coins. This representation of all possible outcomes and their probabilities is known as a probability distribution. Thus, we refer to this as the probability distribution of "number of heads" in the experiment of tossing of three coins. While we see that our previous knowledge on computation of probabilities helps us in arriving at such representations, we recognise that the calculations may be quite tedious. This is apparent, if you try to calculate the probabilities of different number of heads in a tossing of twelve coins. Developments in Probability Theory help us in specifying the probability distribution in such cases with relative ease. The theory also gives certain standard probability distributions and provides the conditions under which they can be applied. We will study the probability distributions and their applications in this and the subsequent unit. The objective of this unit is to look into a type of probability distribution, viz., a discrete probability distribution. Accordingly, after the initial presentation on the basic concepts and definitions, we will discuss as to how discrete probability distributions can be used in decision-making.

Activity A

Suppose you are interested in betting on `tails' in a tossing of four coins. Write down the result of the experiment in terms of the "number of tails" (zero to four) that may occur, with their respective probabilities of occurrence. Elaborate as to how this ma] help you in betting.

…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………

10.2 BASIC CONCEPTS : RANDOM VARIABLE AND PROBABILITY DISTRIBUTION

Before we attempt a formal definition of probability distribution, the concept of 'random variable' which is central to the theme, needs to be elaborated.

In the example given in the Introduction, we have seen that the outcomes of the experiment of a toss of three coins were expressed in terms of the "number of heads" Denoting this "number of heads" by the letter H, we find that in the example, H can assume values of 0, 1, 2 and 3 and corresponding to each value, a probability is associated. This uncertain real variable H, which assumes different numerical values depending on the outcomes of an experiment, and to each of whose values a probability assignment can be made, is known as a random variable. The resulting representation of all the values with their probabilities is termed as the probability distribution of H. It is customary to present the distribution as follows :

Probability Distribution of Number of Heads (H)

H	P(H)
0	0.125
1	0.375
2	0.375
3	0.125

In this case, as we find that H takes only discrete values, the variable H is called a discrete random variable and the resulting distribution is a discrete probability distribution.

In the above situation, we have seen that the random variable takes a limited number of values. There are certain situations where the variable of interest may take infinitely many values. Consider for example that you are interested in ascertaining the probability distribution of the weight of the one kilogram tea pack, that is produced by your company. You have reasons to believe that the packing process is such that the machine produces a certain percentage of the packs slightly below one kilogram and some above one kilogram. It is easy to see that there is essentially to chance that the pack will weigh exactly 1.000000 kg., and there are infinite number of values that the random variable ".weight" can take. In such cases, it makes sense to talk of the probability that the weight will be between two values, rather than the probability of the weight will be between two values, rather than the probability of the weight taking any specific value. These types of random variables which can take an infinitely large number of values are called continuous random variables, and the resulting distribution is called a continuous probability distribution. Sometimes, for the sake of convenience, a discrete situation with a large number of outcomes is approximated by a continuous distribution: Thus, if we find that the demand of a product is a random variable taking values of 1, 2, 3... to 1000, it may be worthwhile to treat it as a continuous variable. Obviously, the representation of the probability distribution for a continuous random variable is quite different from the discrete case that we have seen. We will be discussing this in a later unit when we take up continuous probability distributions.

Coming back to our example on the tossing of three coins, you must have noted the presence of another random variable in the experiment, namely, the number of tails (say T). T has got the same distribution as H. In fact, in the same experiment, it is possible to have some more random variables, with a slight extension of the experiment. Supposing a friend comes and tells you that he will toss 3 coins, and will pay you Rs. 100 for each head and Rs. 200 for each tail that turns up. However, he will allow you this privilege only if you pay him Rs. 500 to start with.

You may like to know whether it is worthwhile to pay him Rs. 500. In this situation, over and above the random variables H and T, we find that the money that yciu may get is also a random variable. Thus,

if H =number of heads in any outcome, then 3 - H = number of tails in any outcome (as the total number of heads and tails that can occur in a toss of three coins is 3)

The money you get in any outcome = 100H + 200 (3 - H)

= 600 -100H = x (say)

We find that x which is a function of the random variable H, is also a random variable.

We can see that the different values x will take in any outcome are

(600 -100 x 0) =600

(600-1010 x 1) =500

(600-100 x 2 =400

(600-100 x 3) =300

Hence the distribution of x is :

X	p(X)
600	$\frac{1}{8}$
500	$\frac{3}{8}$
400	$\frac{3}{8}$
300	$\frac{1}{8}$

The above gives you the probability of your getting different sums of money. This may help you in deciding whether you should utilise this opportunity by paying Rs. 500.

From the discussion on this section, it should be clear by now that a probability distribution is defined only in the context of a random variable or a function of a random variable. Thus in any situation, it is important to identify the relevant random variable and then find the probability distribution to facilitate decision-making.

In the next section we will look at the properties of discrete probability distributions and discuss the methods for finding and assessing such distributions.

Activity B

Suppose three units of a product are tested. The result of the test is given in terms of pass or fail. If the probability that a unit will pass inspection is 0.8, find the probability distribution of the number of units that pass inspection.

……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………
……………………………………………………………………………………………

10.3 DISCRETE PROBABILITY DISTRIBUTIONS

In the previous section we have seen that a representation of all possible values of a discrete random variable together with their probabilities of occurrence is called a discrete probability distribution. The objective of this section is to look into the properties of such distributions, and discuss the methods for assessing them.

In discrete situations, the function that gives the probability of every possible outcome is referred to in Probability Theory as the "probability mass function" (p.m.f.).The outcomes, as you must have noted, are mutually exclusive and collectively exhaustive Thus, a representation of the p.m.f. of the number of heads H, in a toss of three coins can be :

$$f(H) = \left.\begin{array}{l} 0.125 \\ = 0.375 \\ = 0.375 \\ = 0.125 \end{array}\right] \begin{array}{l} \text{if } H = 0 \text{ heads} \\ \text{if } H = 1 \text{ heads} \\ \text{if } H = 2 \text{ heads} \\ \text{if } H = 3 \text{ heads} \end{array}$$

Thus, we see that p.m.f. is the name given to a discrete probability distribution, and if, for any situation, we can specify the p.m.f. of the relevant random variable, the whole probability distribution is then specified. The properties of any p. m. f. , say f(x) where x the random variable, can be derived from the fact that f(x) basically refers to probability values. Any probability measure is by definition non-negative f(x) Moreover, it follows from probability theory, that $\sum f(x) = 1$, the sum being taken over all the possible outcomes.

Sometimes, we are interested in finding the probability of a group of outcomes. In such cases, an addition of the relevant values gives us the result. Thus, in the example given earlier, we find that the probability of 2 or 3 heads = f(2) + f(3) = .5. Further, we may be interested in the probability that the random variable will take values less than or equal to a particular quantity. The result in such situations is achieved by specifying what is known as cumulative distribution function (c.d.f.). The c.d.f. denoted by F(H) is formed by adding the probabilities up to a given quantity, and it gives the probability that the random variable H will take a value less than or equal to that quantity. The F(H) in the example discussed earlier can be written as :

$$F(H) = \begin{cases} 0.125 & \text{for } H = 0 \\ 0.5 & \text{for } H = 1 \text{ or less} \\ 0.875 & \text{for } H = 2 \text{ or less} \\ 1.0 & \text{for } H = 3 \text{ or less} \end{cases}$$

we can see from the above c.d.f. that the probability of getting 2 or less heads is 0.875.

Assessment of the p.m.f. of a random variable follows directly from the different approaches to probability that we have discussed in the earlier unit. The different methods by which p.m.f. of a random variable can be specified are :

1 using standard functions in probability theory
2 using past data on the random variable
3 using subjective assessment.

We now discuss each of the methods and the situations where these can be applied.

Using Standard Functions

Sometimes the knowledge of the underlying process in an experiment helps us to specify the probability mass function. Probability theory has come out with standard functions and the conditions under which these standard functions can be applied to any experiment. Consider again the p.m.f. for the random variable H in the tossing of three coins. An alternative way of specifying f(H) would be as follows :

$$f(H) = \frac{3!}{H!(3-H)!} \left(\tfrac{1}{2}\right)^{H} \left(\tfrac{1}{2}\right)^{3-H}, \text{ for } H = 0, 1, 2, 3.$$

Where H ! is read as H factorial and is given by :

$$H!\, 1\times 2\times 3\times \ldots\ldots\ldots\ldots \times H. \text{ and } 0! = 1$$

Thus, for $H = 0$, we have $f(0) = \frac{3!}{0!3!} \left(\tfrac{1}{2}\right)^0 \left(\tfrac{1}{2}\right)^3 = \tfrac{1}{8}$

Similarly, you can verify that the values you get for f(1), f(2), f(3) by substituting 1, 2 and 3 in the above function, are the same as obtained those obtained earlier.

This form of f(H) is made possible, as the coin tossing experiment satisfies the conditions specific to a **Bernoulli Process.** Bernoulli Process is defined in probability theory as a process marked by dichotomous outcomes with probability of an event remaining constant from trial to trial. In coin tossing, we find that the outcome of any toss is either a head or a tail, so that the dichotomy is preserved. Also in each of three coin tosses, the probability of head (or tail) remains constant, namely $\frac{1}{2}$. The probability distribution pertaining to such a process is standardised in probability theory, so that we can directly write down the p.m.f. corresponding to any experiment that satisfies the Bernoulli Process. Such standard discrete distributions will be discussed in detail in a later section.

Using Past Data

Past data on the variable of interest is used to assess the p.m.f., only if we have reasons to believe that conditions similar to the past will prevail. The frequency of occurrence of each of the values of the variable are noted down and the relative frequency of each of the values is taken as a probability measure. The basis lies in the Relative Frequency Approach discussed in the last unit. You may like to compare the resulting p.m.f. with the corresponding frequency distribution. Thus, under the assumption that buyer behaviour has not changed much, we take the past sales data of a product to find the probability distribution of future sales. While frequency distribution is simply a representation of what has happened in the past, p.m.f. represents what we can expect in the future. If you refer now to Example 4 of the last unit, you can see that the probability distribution of the random variable "daily sales of Indian Express" has been estimated from past data. If we denote the random variable by x, we can write down the p.m.f. as :

$$f(x) = \begin{cases} 73/365 \text{ for } x = 85 \\ 146/365 \text{ for } x = 95 \\ 60/365 \text{ for } x = 105 \\ 86/365 \text{ for } x = 110 \end{cases}$$

This method of assessing the p.m.f. stems from the Subjectivists' Approach to probability. This method is applied if there is no past data, and the situation of interest does not resemble any known processes in Probability theory. Suppose a record manufacturing company is contemplating the introduction of a new **ghazal** singer. ' Before introducing him, they want to find out the likely sales of an L.P. record of the new person in the first year of the release of the record. The random variable here is the "sales in first year". Let us denote it by S. We may here use our subjective assessment to find the p.m.f. of S. One way to assess this may be as follows. The company knows that currently one lakh people buy their records and it believes that out of this one lakh people, 20% i.e. 20,000 customers have the attitude to try anything new, so that the other 80,000 will never buy an unknown singer's record in the first year of release. They have also assessed that at least 10% of their customers are always ready for new ghazals. Building up on such assessments, the final p.m.f. of S may be :

$$f(s) = \begin{cases} .6 \text{ for } S = 10{,}000 \\ .2 \text{ for } S = 15{,}000 \\ .2 \text{ for } S = 20{,}000 \end{cases}$$

In other words, they expect that sales in the first year will be 10,000 with a 60% chance, and 20% chance each that 15,000 or 20,000 people will buy it.

We have seen the different ways to assess a discrete probability distribution. These distributions help us in our decisions by presenting the total scenario in an uncertain situation. The p.m.f. of sales as discussed above, may help the company in deciding how many records should be produced in the first year. While producing 10,000 records is definitely a safe thing to do, we realise that a 40% chance of not being able to meet demand is also there. Similarly production of 20,000 records takes care of meeting all demands that may arise, but then there is a chance that some records may not be sold. Systematic analysis of such decisions can be done with the p.m.f. and the relevant cost data, and will be taken up in Unit 12. Analysis is made easier, if together with the p.m.f. data, certain key figures of the p.m.f. are presented. Thus, it may be easier for us to see things, if the expected sales figure is given to us in the above case. These key figures pertaining to a p.m.f. are called summary measures. In the next section we discuss some summary measures that are helpful in analyzing situations.

Activity C

Cheek whether the following p.m.f. applies for the random variable in activity B

$$f(X) = \frac{3!}{X!(3-X)!}(.8)^{X}(.2)^{3-X}$$

where X = the number of units that pass inspection

(Hint : find f(0), f(1), f(2) and f(3) by substituting X = 0, 1, 2, and 3 in the above function. Check whether these values are the same as what you obtained earlier.)

…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………

10.4 SUMMARY MEASURES AND THEIR APPLICATIONS

As the name implies, a summary measure of a probability distribution basically summarizes the distribution through a single quantity. Just as we have seen in the case of a frequency distribution, here too we have the measure of location and dispersion that help us to have a quick picture of the behaviour of the random variable concerned. The objective of this section is to look into some of the summary measures and discuss the possible application of these measures.

Measures of Location

The most widely used location measure is the Expected Value. It is similar to the concept. of mean of a frequency distribution and is calculated as the weighted average of the values of the random variable, taking the respective probabilities of occurrence as the weight. Thus, in the tossing of three coins, the Expected Value of Number of Heads, written as E(H) can be found as follows :

$$E(H) = \sum H \times f(H) = 0 \text{ x } .125 + 1 \text{ x } .375 + 3 \text{ x } .125 = 1.5$$

Similarly, considering the extension of the experiment as discussed earlier, we can calculate the money you can expect if you take up your friend's proposal, as :

$$E(X) = 600 \text{ x } .125 + 500 \text{ x } .375 + 400 \text{ x } .375 + 300 \text{ x } .125 = \text{Rs. } 450$$

Recalling that you have to pay Rs. 500 to get the privilege of entering this game, you may decide not to go in for it as the expected pay off is less than the sum you have to pay. It may be noted in this context that the pay off X at any outcome is a function of the random variable H. As already noted, X itself is a random variable. Instead of calculating the E(X) as above, it is possible to calculate the E(X) as follows :

$$E(X) = E(600 - 100H) = 600 - 100E(H) = 600 - 100 \text{ x } 1.5 = 450$$

It can be seen that for any linear function g(H) of H, the following holds : E[g(H)] = g[E(H)]. That this is not true, for functions other than linear can be verified by taking, for example, $g(H) = H^2$

$$E(H^2) = \sum H^2 \, f(H) = 0 \text{ x } .125 + 1 \text{ x } .375 + 4 \text{ x } .375 + 9 \text{ x } .125 = 3$$

However $[E\,(H)]^2 = (1.5)^2 = 2.25$

Thus $[E\,(H)]^2$ # $E\,(H^2)$.

Expected value of a random variable gives us a measure of location and is an indicator of the long-run average value that we can expect. In the computation of the expected value, the most likely outcome is given the highest weight age. Sometimes, it is useful to characterize the probability distribution by the most likely value, which is defined as the mode. The modal value is the vat 'e corresponding to which, the probability of occurrence is maximum. Another met Sure of location that is of interest is known as 'fractal'. A value H_z is defined as the k fractal of the distribution of H, if

$$F(H) \leq k \text{ for all } H < H_z$$

$$\text{and } F(H) \geq k \text{ for all } H \leq H_z$$

Recalling the c.d.f. of H, we have developed earlier

H	F(H)	
0	.125	
1	.500	← .60
2	.875	
3	1.000	

Suppose we want to find the .60th fractal of the distribution, i.e., we want to find a value of $H = H_k$ such that $F(H) \leq .60$ for $H < H_k$ and $F(H) \geq .60$ for all $H \geq H_k$. We identify that .60 lies between .50 and .875 F(H) values. This is shown by an arrow in the above distribution. The value of H just above it is one that will be the .60th fractile H = 2 is the required answer. We can verify that for H < 2 i.e. for H = 0 and 1, F (0) = .125 and F(1) = .5, both of which are less than 0.6. Similarly for all $H \geq 2$, F(2) = .875 and F(3) = 1, both of which are greater than .60. Hence it satisfies the conditions.

You may note that the .50th fractile here is 1, i.e. if any required fractile coincides with any F(H) value in the distribution then the value with which it matches, is the required value. You may verify whether this satisfies the stated conditions. The .5th fractile is called the median of the distribution and is of interest at times.

Measures of Dispersion

Standard Deviation (SD), range and absolute deviation are the measures of dispersion of a distribution. Of these, SD being the most widely used, we will discuss it here. You may recall that the same term has been used in the context of a frequency distribution also. However, in a discrete probability distribution, we are dealing with a random variable, and the distribution represents various values of the random variable that we expect will occur in the future. In such, cases, the variance is defined as the expected value of the square of the difference between the random variable and its expected value. Then SD is given by the square root of the variance. Thus, for the random variable H in the coin tossing example, we can write :

$$
\begin{aligned}
\text{Variance} &= E[H - E(H)]^2 = E[H - 1.5]^2 \\
&= (0-1.5)^2 f(0) + (1-1.5)^2 f(1) + (2-1.5)^2 f(2) + (3-1.5)^2 f(3) \\
&= 2.25 \times \frac{1}{8} + 0.25 \times \frac{3}{8} + 0.25 \times \frac{3}{8} + 2.25 \times \frac{1}{8} \\
&= \frac{3}{4}
\end{aligned}
$$

$$\text{and S.D.} = \sqrt{\text{variance}} = \tfrac{1}{2}\sqrt{3}$$

The knowledge on expected value and standard deviation of a distribution of a random variable is useful in our decisions. Suppose you have got an offer to take up any one of the two projects A and B. Both A and B have got uncertain outcomes, so that the payoff for A and B are random variables. If expected payoff for project A is equal to that of project B, and S. D. of payoff in the case of A is less than that of B, then you may decide to choose project A. Here S.D. summarises the variability in monetary payoffs that we can expect from the projects.

We now take up an example to illustrate the use of expected value in decision-making. More complex situations will be taken up later when we study Decision Theory.

Example 1

Consider a newspaper seller who gets newspapers from the local office of the Newspaper every morning and sells them from his shop. He buys each copy for 60 p. and sell it for Rs. 1.10p. However, he has to tell the office in advance as to how many copies he will buy. The office takes back the copies he is not able to sell and pays him only 30 p. for each copy. His problem is essentially to find out how many copies he should order every day. He has estimated the p.m.f. of the daily demand from past data

$$
f(D) = \begin{cases}
0.1 \text{ for } D = 30 \\
0.2 \text{ for } D = 31 \\
0.2 \text{ for } D = 32 \\
0.3 \text{ for } D = 33 \\
0.1 \text{ for } D = 34 \\
0.1 \text{ for } D = 35
\end{cases}
$$

Solution

To analyse such situations, first we formalise the problem in terms of alternative courses of actions open to the newspaper man. As he expects that the daily demand will not be less than 30 or more than 35, we understand that there is no point in his ordering less than 30 or more than 35 copies. Thus, he has got six options :

Alternative 1. Order 30 copies
Alternative 2. Order 31 copies
Alternative 3. Order 32 copies

Alternative 4. Order 33 copies
Alternative 5. Order 34 copies
Alternative 6. Order 35 copies

Corresponding to each alternative action, there are six possible values that the demand can take and each of these values lead to a monetary payoff with different chances of occurrence. We can calculate the expected monetary payoff fat each alternative and choose the alternative that promise us the highest expected payoff.

For calculating monetary payoff corresponding to any outcome and any action, we note:

1 If he orders X copies and demand (D) turns out to be more than or equal to X, then he will be able to sell only X copies, so that the payoff will he (1-10 - 0.60) x X = 0.50 X

2 If he orders X copies and D turns out to be less than X, then he will be able to sell D copies for which he will profit 0.5 D and he will be losing (.60 - .30) = 30 p. for each copy he ordered more, i.e. loss = .30 (X-D).

His payoff = .5D - .3 X + .3D

= .8D - .3X

With the above background, we are now in a position to calculate the payoff P corresponding to each outcome of an alternative. As these payoff values correspond to the demand values only, the chances of occurrence of the payoffs are given by the chances of occurrence of the respective demand figures. Thus, for each alternative, the p.m.f. of P and the corresponding Expected value of P can be calculated. A sample calculation for Alternative 4 (order 33 copies) is shown below.

Alternative 4.

Order 33 copies (X = 33)

Outcome	Demand(D)	If D?X then P=.5 X If D<X then P= .8D - .3X	P	f(P)
1	30	P=.8x30-.3x33	14.1	.1
2	31	P=.8x31-.3x33	14.9	.2
3	32	P = .8 x32 - .3x33	15.7	.2
4	33	P=.5x33	16.5	.3
5	34.	P=.5x33	16.5	.1
6	35	P=.5 x33	16.5	.1

E(P) = 14.1 x .1 + 14.9 x .2 + 15.7 x .2 + 16.5 x .3 + 16.5 x .1 + 16.5 x .1 = 1.41 + 2.98 + 3.14 +4.95 + 1.65 + 1.65 = 15.78

Similarly, we can calculate the Expected payoff for other alternatives also. The newspaper man should go for the alternative that gives him the highest expected payoff A convenient representation of the alternatives and the outcomes is given below. Corresponding to alternative 4, we have filled up the values. You may now fill up the other cells.

Probabilities of Demand	.1	.2	.2	.3	.1	.1	Expected Payoff E(P)
Demand (Outcomes) Order (Alternative)	30	31	32	33	34	35	
1. 30							
2. 31							
3. 32							
4. 33	14.1	14.9	15.7	16.5	16.5	16.5	15.78
5. 34							
6. 35							

On solving E(P), we find that the maximum expected payoff is obtained for Alternative 4. Hence we can say that the newspaper man should order for 33 copies.

Activity D

In the above problem, instead of calculating the payoffs, we could have calculated the expected opportunity loss for each alternative.

We recognise that for each alternative and an outcome, three situations can arise:

1 Number ordered (X) = Number demanded (D) : In this case there is no loss to the newspaper man as he has stocked the right number of copies.

2 Number ordered (X) < Number demanded (D) : In this case, he has understocked. and for each copy that he has not ordered for and could have sold, he loses the profit = 0.50 p. Thus, opportunity loss = .50 (D-X).

3 Number ordered (X) > Number demanded (D) : In this case he has ordered for more than he can sell, so he loses (.60-.30) = .30 p. for each extra copy that he has ordered therefore opportunity loss = 0.30 (X-D).

Using the above, calculate the opportunity loss corresponding to each outcome of each alternative. Find the Expected opportunity loss for each alternative and state how you will decide on the basis of these expected values.

…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………
…………………………………………………………………………………………

10.5 SOME IMPORTANT DISCRETE PROBABILITY DISTRIBUTIONS

While examining the different ways of assessing p.m.f., we have noted that proper identification of experiments with certain known processes in Probability theory helps us in writing down the probability distribution function. Two such processes are the Bernoulli and the Poisson. The standard discrete probability distribution that are consequent to these processes are the Binomial and the Poisson distribution. The objective of this final section is to look into the conditions that characterise these processes, and examine the standard distributions associated with the processes. This will enable us to identify situations for which these distributions apply.

Bernoulli Process

Any uncertain situation or experiment that is marked by the following three properties is known as a Bernoulli Process.

1 There are only two mutually exclusive and collectively exhaustive outcome[s] ' ^ the experiment..

2 In repeated observations of the experiment, the probabilities of occurrence of these events remain constant.

3 The observations are independent of one another.

Typical examples of Bernoulli process are coin-tossing and success-failure situations. In repeated tossing of coins, for each toss, there are two mutually exclusive and collectively exhaustive events, namely, head and tail. We also know that the probability of a head or a tail remains constant (= $\frac{1}{2}$)from toss to toss, and result of one toss does not effect the result of any other toss.

Similar dichotomy is preserved in testing of different pieces of a product. Each piece when tested may be defective (a failure) or non-defective (a success). We know that the production process is such that the probability of a non-defective in any trial is P and that of a defective = q = (1 - p)

Once the process has stabilised, it is reasonable to assume that the success and failure of each piece is independent of the other and also the probability of a success (p) or a failure (q) remains constant from trial to trial. Thus, it satisfies the conditions of a Bernoulli process.

The random variables that may be of interest in the above situations are :

1 The number of successor failure in a specified number of trials, given the knowledge on the probability of a success in trial. This implies that if the experiment is observed n times then given that the probability of a success is fin any observation, we are interested in finding out the distribution of number of successes that may occur in n observations.

2 The number of trials needed to have a specified number of successes, given the knowledge on the probability of success in any trial. We are interested in finding out the probability distribution of the number of trials required to get a specified number of successes.

The Binomial distribution and the Pascal distribution provide us with the required p.m.fs. in the above two cases. We discuss these two distributions with examples.

Binomial Distribution

Let us take the example of a machining process which produces on an average 80% good pieces. We are interested in finding out the p.m.f. of the number of good pieces in 5 units produced from this process. From our definition, this situation is a Bernoulli process, with the probability of success = P = 0.8

:. Probability of failure or defective pieces = q =1 - P = 0.2.

The number of trials = 5.

Let n be the random variables of interest, i.e. the number of good pieces. As N = 5, obviously r can take values of 0, 1, 2, 3, 4, 5, i.e. as 5 pieces are produced, at the best all 5 can be good pieces. We can now try to calculate the probabilities for different values of r using the results given in the last unit :

r = 0 means all 5 are failure. As the probability of failure is q in every trial, and the trials ,We independent, probability of 5 failures = q x q x q x q x q = q^5. The total number of outcomes in the experiment are 2^5 and we find that only in one outcome all 5 are failures.

Therefore $f(0) = q^5$

r = 1 implies that there is one success and four failures. The probability of this is pq^4 However, out of the 2^5 possible outcomes, one success and four failures can occur in the following ways :

1st unit is a success and the rest are failure i.e. SFFFF

2nd unit is a success and the rest are failure i.e. FSFFF

3rd unit is a success and the rest are failure i.e. FFSFF

4th unit is a success and the rest are failure i.e. FFFSF

5th unit is a success and the rest are failure i.e. FFFFS

where S denotes a success and F a failure. Thus, 1 success and 4 failures can occur in 5 different ways, for each of which the probability is pq4

Hence $f(1) = 5\,pq^4$. Similarly for r = 2, the probability of 2 successes and 3 failures is $p^2\,q^3$**.** To find the number of outcomes in which 2S and 3F will occur we can use the following. Basically, we want to know the different ways in which 2S and 3F can be put in a sequence. This is represented by ${}^5C_2$ read as "five C two" and given by

$$\frac{5!}{3!2!} = 10$$

Hence $f(2) = 10p^2q^3$

The required p.m.f. of r is then

$$f(r) = \begin{cases} q^5 \text{ for } r = 0 \\ 5pq^4 \text{ for } r = 1 \\ 10p^2q^3 \text{ for } r = 2 \\ 10p^3q^2 \text{ for } r = 3 \\ 5\,p^4q \text{ for } r = 4 \\ p^5 \text{ for } r = 5 \end{cases}$$

Each of the terms for r = 0 .... 5 correspond to the binomial expansion of $(q + p)^5 = q^5 + 5pq^4 + 10p^2q^3 + 10p^3q^2 + 5p^4q + p^5$, hence the above distribution is known as Binomial distribution.

In general, as Binomial distribution gives the probability of r successes in n trials as

$$f(r) = {}^nC_r\, p^r q^{n-r}$$

$$\text{where, } {}^nC_r = \frac{n!}{r!\,(n-r)!}$$

p = probability of success in any trial

q = probability of failure in any trial = 1-p.

often f(r) is written as f(r/n, p), as n and p are given.

We can verify that the above has got the properties of a p.m.f. We can write down directly the p.m.f. as above for any situation that satisfies the earlier stated conditions.

Given the standard expression, it is possible to calculate the expected value (referred to as the mean) and the variance of a Binomial distribution :

$$\text{Expected value (Mean)} = \Sigma r f(r) = \Sigma r \cdot \frac{n!}{r!\,n-r!} p^r q^{n-r}$$

$$\text{As } n! = n \times n-1 \times n-2 \;\ldots\ldots \times 1 = n(n-1)!$$

$$r! = r(r-1)\,!$$

and $\Sigma \dfrac{(n-1)!}{(r-1)!\,(n-r)!} p^{r-1} q^{n-r} = 1$ being sum of probabilities of all outcomes of number of successes in n − 1 trials.

$$\text{Mean } = \Sigma n \frac{(n-1)!}{(r-1)!\,(n-r)!} p.p^{r-1}q^{n-r}$$

$$= np\,\Sigma \frac{(n-1)!}{(r-1)!\,(n-r)!} p^{r-1}q^{n-r} = np$$

The variance of the distribution can be shown to be npq.

As, n, p, q, are given constants for a particular distribution, the mean and variance are also constant. These are called parameters of a distribution and are often used to specify a distribution.

Pascal Distribution

Suppose we are interested in finding the p.m.f. of the number of trials (n) required to get 5 successes, given the probability p, of success in any trial.

We see that 5 successes can be obtained only in 5 or more trials. Thus, we want to find f(n) for n = 5, 6……………….etc.

If n trials are required to get 5 successes then the last trial has to result in a success, while in the rest of the n-1 trials, 4 successes have been obtained. This implies that :

f(n) = (probability of 4 successes in n-1 trials) X p.

$= {}^{n-1}C_4\, p^4 q^{n-5}\, .p$

It is customary to write f(n) as f(n/r, p), as r and p are given here. The above satisfies the properties of a p.m.f. The mean and the variance of the distribution are $\frac{r}{p}$ and $\frac{rq}{p^2}$ respectively.

Of the many standard discrete distributions, we have so far discussed the Binomial and the Pascal. We now present the Poisson distribution which is applicable to events occurring randomly over time and space. This p.m.f. has been used widely to represent distributions of several random variables like demand for spare parts, number of telephone calls per hour, number of defects per metre in a bale of cloth, etc. In order to apply this p.m.f. in any situation, the conditions of a Poisson process need to be satisfied. We discussed these conditions and the Poisson distribution in the following paragraphs.

Poisson Process and Poisson Distribution

Conditions specific to the Poisson process are easily seen by establishing them in the context of the Bernoulli process. Let us consider a Bernoulli process with n trials and the probability of success in any trial = $\frac{m}{n}$, where $m \geq 0$. Then we do now that the probability of r successes in n trials is given by:

$$f(r) = {}^nC_r \left(\frac{m}{n}\right)^r \left(1 - \frac{m}{n}\right)^{n-r}$$

We note that $${}^nC_r = \frac{n!}{r!(n-r)!} = \frac{n(n-1)(n-2)(n-r+1)(n-r)!}{r!(n-r)!}$$

$$f(r) = \frac{m^r}{r!}\frac{n}{n} \times \frac{n-1}{n} \times \frac{n-2}{n} \times \ldots\ldots \frac{(n-r+1)}{n} \times \left(1 - \frac{m}{n}\right)^{-r} \times \left(1 - \frac{m}{n}\right)^n$$

Now, if $n \to \infty$, then the terms, $\frac{n-1}{n}$, $\frac{n-2}{n}$, ... $\frac{n-r+1}{n}$

and $\left(1 - \frac{m}{n}\right)^{-r}$ will all be tending to 1.

Also from a theorem in Calculus, it is known that $\left(1 - \frac{m}{n}\right)^n$

tends to e^{-m}, if $n \to \infty$. Thus, we have $f(r) = \frac{e^{-m} \times m^r}{r!}$

The above function is a Piosson p.m.f. Thus, a Poisson process corresponds to a Bernoulli process with a very large number of trials (n) and with a very low probability of success (m/n) in any trial. We will now demonstrate a real life analogy of such a process.

Consider the occurrence of any uncertain event over time or space in such a way that the average occurrence of the event over unit time or space is m. We may take the number of accidents occurring over a time period with m denoting the average number of accidents per month; or we may be interested in the number of defects occurring in a strip if cloth manufactured by a mill, with m denoting the average number of defects per metre. For each of such situations, we see the possibility of dividing the time or space interval into n very small segments such that within a small segment the conditions of the Bernoulli process hold. Thus, one month can be divided into (say) 30 x 24 x 60 intervals of one minute each, so that the probability of occurrence of an accident in any

minute = $\frac{m}{30 \times 24 \times 60}$, and reduces to a very small quantity, so that there is almost no chance of having two accidents occurring in one minute, The independence property of the Bernoulli trial also holds true here, as a one minute interval basically corresponds to a trial. Similar possibilities also exist in the cloth example.

The above enables us to calculate the probability that r accidents will occur, from the Poisson formula derived earlier. As we have made n very large, and p very small, and have also verified that the Bernoulli conditions are satisfied, we can write f(r) = $\frac{e^{-m}m^r}{r!}$ as the required p.m.f. in such a cases.

The p.m.f. is alternatively written as f(r/m).

Suppose we want to find the distribution of the number of accidents r, given that there are, on an average, 3 accidents per month. We can find this by putting r = 0, 1, 2, 3, 4,…………………in f(r/3)

$$f(0/3) = \frac{e^{-3} \times 3^0}{O!} = e^{-3} = .0498.$$

The mean and variance of a Poisson distribution are equal and are given by m. This property is sometimes used to check whether the Poisson applies for the event under study.

Activity E

A plane has got 4 engines. The probability of an engine failing is 1/3 and each engine may fail independently of the other engine. Find the probability that all the engines will fail. Write down the p.m.f. of 'Failed Engines'

…………………………………………………………………………………………
…………………………………………………………………………………………

………………………………………………………………………………………………
………………………………………………………………………………………………

Activity F

If 1% of the bolts produced by a certain machine are defective, find the probability that in a random sample of 300 bolts, all bolts are good.

[Hint : This is a case of a Binomial distribution with n = 300 and p = .01. We have to find f (0/300, .01). As n is large (300) and p is small (.01), Poisson can be used to calculate the required probability. Poisson with m = np = 300 *x* .01 = 3 will lead to the answer, i.e., find f(0/3).]

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

Activity G

From past experience a Proof reader has found that after he proofreads, there remain 2 errors on an average in a page. What is the probability of finding a page without any error?

………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………
………………………………………………………………………………………………

10.6 SUMMARY

We have introduced the concepts of random variable and probability distribution in this unit. In any uncertain situation, we are often interested in the behaviour of certain quantities that take different values in different outcomes of the experiments. These quantities are called random variables and a representation that specifies the possible values a random variable can take, together with the associated probabilities, is called a probability distribution, The distribution of a discrete variable is called a discrete probability distribution and the function that specifies a discrete distribution is called a probability mass function (p.m.f.). We have looked into situations that gives rise to discrete probability distributions, and discussed how these distributions are helpful in decision-making. The concept and application of expected value and other summary measures for such distributions have been presented. Different methods for assessing such distributions have also been discussed. In the final section certain standard discrete probability distributions and their applications have been discussed.

10.7 FURTHER READINGS

Gangolli, R.A. and D. Ylvisaker, *Discrete Probability,* Harcourt, Brace & World, Inc.: New York.

Levin, R.I., 1984. *Statistics for management,* Prentice-Hall, Inc. : Englewood-Cliffs.

Parzen,E., 1960. *Modern Probability Theory and its Applications,* Wiley: New York.