6. Discrete Probability Distributions

1. *Probability Distributions*

This chapter explains the concepts and applications of what is called a probability distribution. In addition, special probability distributions, such as the binomial, multinomial, Poisson, and hyper-geometric distributions, are explained.

- **Random variable** is a variable whose values are determined by chance.
- **Discrete variables** which have a finite number of possible values or an infinite number of values that can be counted. The word counted means that they can be enumerated using the numbers 1, 2, 3, etc. For example, the number of family members (1, 2, 3, 4, …), number of calls, and so on.

Example 1: three coins are tossed, the sample space is represented as { TTT, TTH, THT, HTT, HHT, HTH, THH, HHH}; if X is the random variable for the number of heads, then X assumes the value 0, 1, 2, or 3. (X = 0, 1, 2, 3)

(X: 0 = no head, 1 = one head, 2 = two head, 3 = three head)

Probabilities for the values of *X* can be determined as follows:

No heads	One head			Two heads			Three heads
TTT	TTH	THT	HTT	HHT	HTH	THH	HHH
$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
$\frac{1}{8}$	$\frac{3}{8}$			$\frac{3}{8}$			$\frac{1}{8}$

Number of heads X	0	1	2	3
Probability $P(X)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

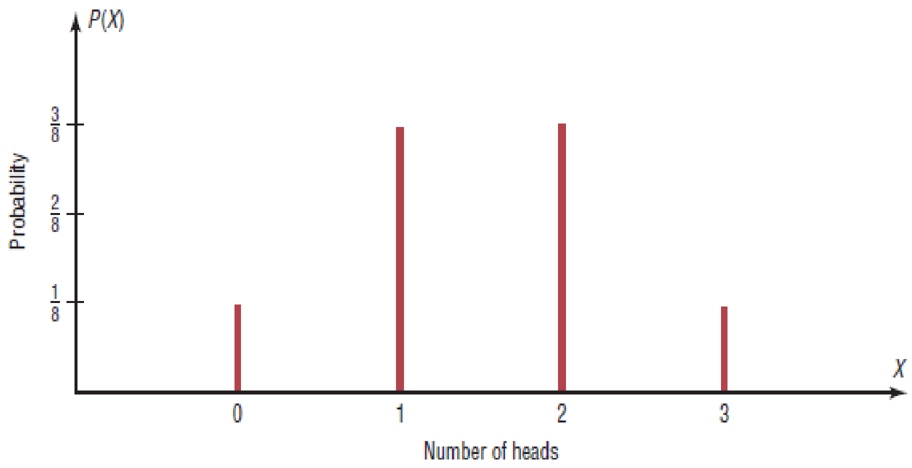
- A **discrete probability distribution** consists of the values a random variable can assume and the corresponding probabilities of the values. The probabilities are determined theoretically or by observation.
- **Discrete probability distributions** can be shown by using a graph or a table. Probability distributions can also be represented by a formula

Example 2: Represent graphically the probability distribution for Example 1.

Number of heads X	0	1	2	3
Probability $P(X)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Solution :

The values that X assumes are located on the x axis, and the values for $P(X)$ are located on the y axis.



Note: Two Requirements for a Probability Distribution

1. The sum of the probabilities of all the events in the sample space must equal 1; that is, $\sum \boldsymbol{P(X) = 1}$.
2. The probability of each event in the sample space must be between or equal to 0 and 1. That is, $\boldsymbol{0 \leq P(X) \leq 1}$.

Example 3: Determine whether each distribution is a probability distribution.

a.

X	4	6	8	10
$P(X)$	−0.6	0.2	0.7	1.5

b.

X	1	2	3	4
$P(X)$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$

c.

X	8	9	12
$P(X)$	$\frac{2}{3}$	$\frac{1}{6}$	$\frac{1}{6}$

d.

X	1	3	5	7	9
$P(X)$	0.3	0.1	0.2	0.4	−0.7

Solution :

a) No. It is not a probability distribution since P(X) cannot be negative or greater than 1. **b)** Yes. It is a probability distribution.

c) Yes. It is a probability distribution. **d)** No, since $P(X) \neq -0.7$.

6.2. ***Mean, Variance, and Standard Deviation***

The mean, variance, and standard deviation for a probability distribution are computed differently from the mean, variance, and standard deviation for samples.

A- Mean

The mean of a random variable with a discrete probability distribution is:

$$\mu = X_1 \cdot P(X_1) + X_2 \cdot P(X_2) + X_3 \cdot P(X_3) + \ldots\ldots + X_n \cdot P(X_n)$$

$$= \Sigma X \cdot P(X)$$

Where: $X_1, X_2, X_3, \ldots, X_n$ are the outcomes and $P(X_1), P(X_2), P(X_3), \ldots, P(X_n)$ are the corresponding probabilities.

Note: $\mathbf{\Sigma X \cdot P(X)}$ means to sum the products.

Example 4: Find the mean of the number of spots that appear when a die is tossed.

Solution:-

In the toss of a die, sample space is 1, 2, 3, 4, 5, 6; the mean can be computed thus.

Outcome X	1	2	3	4	5	6
Probability $P(X)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$\boldsymbol{\mu = \Sigma X \,.\, P(X)} = 1.\frac{1}{6} + 2.\ \frac{1}{6} + 3.\frac{1}{6} + 4.\frac{1}{6} + 5.\frac{1}{6} + 6.\frac{1}{6}\ = \boldsymbol{\frac{21}{6} = 3.5}$

Example 5: In a family with two children, find the mean of the number of children who will be girls.

Number of girls X	0	1	2
Probability $P(X)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Solution

The probability distribution is as follows:

The mean is : $\mu = \Sigma X \,.\, P(X) = 0.\frac{1}{4} + 1.\ \frac{1}{2} + 2\ = 1$

B. Variance and Standard Deviation

- For a probability distribution, the mean of the random variable describes the measure of the so-called long-run or theoretical average, but it does not tell anything about the spread of the distribution. To measure this spread or variability, statisticians, the variance and standard deviation are used for this purpose.
- The formula for the variance of a probability distribution is:

$$\sigma^2 = \sum [X^2 . P(X)] - \mu^2$$

- The standard deviation of a probability distribution is:

$$\sigma = \sqrt{\sigma^2} \quad \text{Or} \quad \sqrt{\sum [X^2 . P(X)] - \mu^2}$$

Example 6: A box contains 5 balls. Two are numbered 3, one is numbered 4, and two are numbered 5. The balls are mixed, and one is selected at random. After a ball is selected, its number is recorded. Then it is replaced. If the experiment is repeated many times, find the variance and standard deviation of the numbers on the balls.

Number on ball X	3	4	5
Probability $P(X)$	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{2}{5}$

Solution

Let X be the number on each ball. The probability distribution is

The mean is : $\mu = \Sigma X \,.\, P(X) = 3.\frac{2}{5} + 4.\ \frac{1}{5} + 5\frac{2}{5} = 4$

The variance is

$$\begin{aligned}
\sigma &= \Sigma[X^2 \cdot P(X)] - \mu^2 \\
&= 3^2 \cdot \tfrac{2}{5} + 4^2 \cdot \tfrac{1}{5} + 5^2 \cdot \tfrac{2}{5} = 4 \\
&= 16\tfrac{4}{5} - 16 \\
&= \tfrac{4}{5}
\end{aligned}$$

The standard deviation is

$$\sigma = \sqrt{\frac{4}{5}} = \sqrt{0.8} = 0.894$$

The mean, variance, and standard deviation can also be found by using vertical columns, as shown.

X	$P(X)$	$X \cdot P(X)$	$X^2 \cdot P(X)$
3	0.4	1.2	3.6
4	0.2	0.8	3.2
5	0.4	2.0	10
		$\Sigma X \cdot P(X) = 4.0$	16.8

Find the mean by summing the $\Sigma X \cdot P(X)$ column, and find the variance by summing the $X^2 \cdot P(X)$ column and subtracting the square of the mean.

$$\sigma^2 = 16.8 - 4^2 = 16.8 - 16 = 0.8$$

and

$$\sigma = \sqrt{0.8} = 0.894$$

Example 7: A talk radio station has four telephone lines. If the host is unable to talk (i.e., during a commercial) or is talking to a person, the other callers are placed on hold. When all lines are in use, others who are trying to call in get a busy signal. The probability that 0, 1, 2, 3, or 4 people will get through is shown in the distribution. Find the variance and standard deviation for the distribution.

X	0	1	2	3	4
$P(X)$	0.18	0.34	0.23	0.21	0.04

The mean is

$$\mu = \Sigma X \,.\, P(X) = 1.6$$

The variance is

$$\sigma^2 = \sum [X^2 . P(X)] - \mu^2 = 1.23$$

The standard deviation is:

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.23} = 1.1$$

6.3. The Binomial Distribution (BD):

The Binomial Distribution (BD) is a probability experiment that satisfies the following four requirements:

1. There must be a fixed number of trials.

2. Each trial can have only two outcomes or outcomes that can be reduced to two outcomes. These outcomes can be considered as either success or failure.

3. The outcomes of each trial must be independent of one another.

4. The probability of a success must remain the same for each trial.

6.3.1.Binomial Probability Formula

In a binomial experiment, the probability of exactly X successes in n trials is:

$$\boldsymbol{P(X) = nCx\, P^{x}.\, q^{n-X} = \frac{n!}{(n-X)!\,X!}.\, P^{X}.\, q^{n-X}}$$

$P(S)$ The symbol for the probability of success.

$P(F)$ The symbol for the probability of failure.

p the numerical probability of a success.

q the numerical probability of a failure.

$P(S) = p$ and $P(F) = 1 - p = q$; n The number of trials; X The number of successes in n trials; Note that 0 X n and X 0, 1, 2, 3, . . . , n.

Example 8: A coin is tossed 3 times. Find the probability of getting exactly two heads.

Solution

This problem can be solved by looking at the sample space. There are three ways to get two heads.

HHH, <u>HHT, HTH, THH,</u> TTH, THT, HTT, TTT

The answer is: 1/8+1/8+1/8 = 3/8 = 0.375

In this case, n = 3, X = 2, $p = q = 1/2$

$$\boldsymbol{P(X) = \frac{n!}{(n-X)!\,X!}\,.\,P^X.\,q^{n-X}}$$

$$\boldsymbol{= 0.375 P(X) = \frac{3!}{(3-2)!2!}\,.\,(\frac{1}{2})^2.\,(\frac{1}{2})^1}$$

Example 9: A survey found that one out of five Americans say he or she has visited a doctor in any given month. If 10 people are selected at random, find the probability that exactly 3 will have visited a doctor last month.

Solution

In this case, n = 10, X = 3, p = *1/5*, and q = *4/5*.

$$P(3) = \frac{10!}{(10-3)!3!}\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^7 = 0.201$$

Example 10: A survey from Teenage Research Unlimited (Northbrook, Illinois) found that 30% of teenage consumers receive their spending money from part-time jobs. If 5 teenagers are selected at random, find the probability that at least 3 of them will have part-time jobs.

Solution

To find the probability that at least 3 have part-time jobs, it is necessary to find the individual probabilities for 3, or 4, or 5 and then add them to get the total probability.

$$P(3) = \frac{5!}{(5-3)!3!}(0.3)^3(0.7)^2 = 0.132$$

$$P(4) = \frac{5!}{(5-4)!4!}(0.3)^4(0.7)^1 = 0.028$$

$$P(5) = \frac{5!}{(5-5)!5!}(0.3)^5(0.7)^0 = 0.002$$

Hence,

P(at least three teenagers have part-time jobs)
$= 0.132 + 0.028 + 0.002 = 0.162$

6.3.2.Mean, Variance, and Standard Deviation for the Binomial Distribution

The mean, variance, and standard deviation of a variable that has the *binomial distribution* can be found by using the following formulas.

Mean: $\mu = n \, . \, p$ Variance: $\sigma^2 = n \, . \, p \, . \, q$ Standard deviation: $\sigma = \sqrt{npq}$

Example 11: A die is rolled 480 times. Find the mean, variance, and standard deviation of the number of 3 that will be rolled.

Solution

This is a binomial experiment since getting a 3 is a success and not getting a 3 is considered a failure.

n = 480, p = 1/6, q= 5/6.

Mean: $\mu = n \, . \, P = 480 \times 1/6 = 80$

Variance: $\sigma^2 = n \, . \, p \, . \, q = 480 \times 1/6 \times 5/6 = 66.67$

Standard deviation: $\sigma = \sqrt{npq} = 8.16$

6.4. Other Types of Distributions:

- **The Poisson Distribution**

The probability of X occurrences in an interval of time, volume, area, etc., for a variable where λ (Greek letter lambda) is the mean number of occurrences per unit (time, volume, area, etc.) is:

$$P(X;\ \lambda) = \frac{e^{-\lambda}\lambda^X}{X!}$$

where:

X=0, 1, 2, …; e = 2.718

Example 12: If there are 200 typographical errors randomly distributed in a 500-page manuscript, find the probability that a given page contains exactly 3 errors.

Solution

First, find the mean number λ of errors. Since there are 200 errors distributed over 500 pages, each page has an average of:

$$\lambda = \frac{200}{500} = \frac{2}{5} = 0.4$$

or 0.4 error per page. Since *X* _ 3, substituting into the formula yields

$$P(X;\lambda) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{(2.7183)^{-0.4}(0.4)^{3}}{3!} = 0.0072$$

Example 12: A factory of steel molds, 15 samples have been selected for testing, 10% of the molds have been failed in the test. Determine: (1) more than 2 molds are passed the test. (2) two or less are failed; (3) two are passed.

n= 15; q = 0.1, p = 0.9

$$\boldsymbol{P(X) = \frac{n!}{(n-X)!\,X!}.P^{X}.q^{n-X}}$$

1) P (y ≥ 2) = 1- P (y < 2) = 1- (P (y=1) +P (y=0))

   = 1-[15C0 (0.1)0 (0.9)15 + 15C1 (0.1)1 (0.9)1]

2) P (y ≤ 2) = P (y=2) + P (y=1) + P (y=0) $=15C0\ (0.1)^{0}\ (0.9)^{15} + 15C1\ (0.1)^{1}\ (0.9)^{14} + 15C2\ (0.1)^{2}\ (0.9)^{13} =$

3) $15C2\ (0.9)^{2}\ (0.1)^{13}$