4

Special Probability Distributions

We must be careful not to confuse data with the abstractions we use to analyse them.

4.1 INTRODUCTION

In the previous two chapters we have discussed the probability distribution (discrete and continuous) and their characteristics such as mean (μ) and variance $\left(\sigma^2\right)$. Now, we will discuss some common distributions in which variables are distributed according to some probability law that is in a particular form of pmf (in the case of discrete distribution) or pdf (in the case of continuous distributions). Also, such probability laws are expressed as a mathematical function which are indexed by a quantity called a *parameter*. To understand the word parameter, consider, for example, $f(x; m) = mx,\ x > 0$ represents family of straight lines passing through origin with m is the parameter, and m ranges over real numbers. For a particular value of m (say $m = 1$), we get a member of this family of lines $(y = x)$. Here, in this chapter we shall list some of the most common distributions that arise in applications. In that sense, we call them as *special probability distributions*.

4.2 BERNOULLI DISTRIBUTION

A random variable X which takes two values 0 and 1 with probabilities q and p respectively is called a Bernoulli variate and is said to have a Bernoulli distribution.

These two possible values 0 and 1 of X can be thought of as the only possible outcomes "failure" and "success" of an experiment. Therefore, $q = 1 - p$ or $p + q = 1$. For example, getting a head with a balanced coin (equivalently getting a tail) passing (or failing an examination, getting a defective (or non-defective) item from a lot are certain Bernoulli successes. We refer to an experiment to which the Bernoulli distribution applies as a Bernoulli trial or simply a trial and to sequences of such experiments as repeated trials.

Definition

A random variable X is defined to have a Bernoulli distribution if the probability distribution of X is given by

$$f(x) = \begin{cases} p^x(1-p)^{1-x} & x = 0, 1 \\ 0 & \text{elsewhere} \end{cases}$$

where p is the parameter which satisfies $0 \leq p \leq 1$.

Here $1 - p$ is often denoted by q so that $p + q = 1$.

In our earlier sections we often refer a random experiment tossing 3 coins simultaneously. If we extend this as a repeated trial of say 12 flips of a coin and if we want to know the probability of getting 5 heads, then our earlier approach is really cumbersome. However, we can observe that the probability of getting a head is the same for each of the 12 trials ($p = 1/2$) and there is independence in getting the outcomes (Head or Tail). Hence, if we define "success" as getting a head, under the stated conditions we are interested in finding 5 success in 12 flips or 12 independent repeated trials.

Such repeated trials play an important role in probability and statistics. In such cases, we assume the following.

1. The number of trials is fixed.
2. The parameter p (the probability of success) is the same for each trial.
3. All the trials are independent.

There are certain random variables that arise in connection with repeated trials. One such distribution which concerns the total number of successes (5 heads in 12 flips i.e., number of successes is 5) is a Binomial distribution.

Definition

A random variable X has a **binomial distribution** if the probability distribution of X is

$$f(x) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x} & x = 0, 1, 2 ... n \\ 0 & \text{elsewhere} \end{cases}$$

Here n and p are parameters. We denote $X \sim b(n, p)$ to convey that X has a binomial distribution with parameters n and p. The idea behind this probability distribution is quite natural and we find this idea now. Assume that there are n points (n trials) in a straight line in which x points are called as successes and remaining $n - x$ points are called failures. Using combinatorics we can select $\begin{pmatrix} n \\ x \end{pmatrix}$ ways of 'x' successes. Once the success points are selected, let us apply the probability idea.

Assuming the independence and constant probability (p) of success. we have the probability of getting x successes (and equivalently $n - x$ failures) is $p^x(1 - p)^{n-x}$ or $p^x q^{n-x}$. Hence, the probability of getting x successes in n trials is $\begin{pmatrix} n \\ x \end{pmatrix} p^x q^{n-x}$.

Next we consider the following instance. A sequence of questions is given to a student. As soon as he records his 5^{th} correct answer (5 successes) the test is over and the student is declared as pass. Here we are concerning the number of trials on which the 5^{th} success occurs. Let us assume that answering the correct one is success event and 'p' be the probability of success. Note that the sequence must have a set of 5 questions. (so trials must be 5). Following Table 4.1 explains the situation.

Table 4.1

No. of questions	Answering pattern	Events pattern	Combinatorics and resultant probability
5	All five are correct		p^5
6	1 failure, 5 are correct	Among first 5 questions 4 are success 1 is failure 6^{th} outcome is 5^{th} success	$\binom{5}{4} p^5 q$
7	2 failure 5 successes	First 6 questions have 4 successes 2 failures 7^{th} outcome is 5^{th} success	$\binom{6}{4} p^5 q^2$ and so on.

Note that last column of this table assumes independence of trials and constant probability (p) of success. Extending the idea in the last column, if we have n questions there must be $4(5-1)$ successes and $n-5(=x)$ failures in the first $n-1$ questions and n^{th} outcome must be a success (which is of course 5^{th} success). Hence, the last column can be generalised as $\binom{n-1}{4} p^5 q^{n-5}$ or $\binom{x+5-1}{4} p^5 q^x$. If we generalise this idea of finding the number of trials on which the r^{th} success occurs, in connection with repeated Bernoulli trials, we define a random variable X which is said to have a negative binomial distribution.

Definition

A random variable X has a **negative binomial distribution** if the probability distribution of X is

$$f(x) = \begin{cases} \binom{x+r-1}{r-1} p^r q^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

We denote X as $X \sim nb\,(x_i; r, p)$.

Hence negative binomial distribution represents how long one has to wait for the r^{th} success, so that this random variable is offen referred to as a discrete waiting-time random variable.

If in the negative binomial distribution $r = 1$, then we get a special name for the random variable X. It is called the geometric distribution.

Definition

A random variable X has a **geometric distribution** if the probability distribution of X is

$$f(x) = \begin{cases} pq^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Here p is the only parameter and we denote it as $X \sim Ge(p)$.

One more distribution that arises from repeated trials is Poisson distribution. Specifically, when $n \to \infty$ and $p \to 0$ (equivalently np is constant), the number of successes is a random variable having a poisson distribution. Hence, Poisson distribution is a limiting form of the binomial distribution, which we prove in the following result.

Binomial by Poisson

Out assumption is n is very large and p is very small so that their product np remains constant.

Let $\lambda = np$. Consider $X \sim b(n, p)$, so that

$$\begin{aligned} f(x) &= \binom{n}{x} p^x q^{n-x} \\ &= \binom{n}{x} p^x (1-p)^{n-x} \; x = 0, 1, 2, \cdots, n \end{aligned}$$

Consider $\binom{n}{x} p^x (1-p)^{n-x}$

$$\begin{aligned} &= \frac{n!}{x!\,(n-x)!} p^x (1-p)^{n-x} \\ &= \frac{n(n-1)(n-2)\cdots(n-x+1)}{x!} \frac{\lambda^x}{n^x} \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{1}{x!} \frac{n}{n} \left(\frac{n-1}{n}\right) \left(\frac{n-2}{n}\right) \cdots \left(\frac{n-x+1}{n}\right) \lambda^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{1}{x!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right) \lambda^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\ &= \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{x!} \lambda^x \left(1 - \frac{\lambda}{n}\right)^{n} \cdot \left(1 - \frac{\lambda}{n}\right)^{-x} \\ &= \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)}{x!} \lambda^x \left(1 - \frac{\lambda}{n}\right)^{-x} \cdot \left(1 - \frac{\lambda}{n}\right)^{n} \\ &= 1\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right) \cdots \left(1 - \frac{x-1}{n}\right)\left(1 - \frac{\lambda}{n}\right)^{-x} \cdot \frac{\lambda^x}{x!} \left(1 - \frac{\lambda}{n}\right)^{n} \end{aligned}$$

As $n \to \infty$

$$\begin{aligned} \lim_{n\to\infty} f(x) &= 1 \cdot \frac{\lambda^x}{x!} e^{-\lambda} \qquad x = 0, 1, 2, \cdots \\ &= e^{-\lambda} \left(\frac{\lambda^x}{x!}\right) \quad \text{for} \quad x = 0, 1, 2, \cdots \end{aligned}$$

Definition

A random variable X has a **poisson distribution** if the probability distribution of X is

$$f(x) = \begin{cases} e^{-\lambda}\dfrac{\lambda^x}{x!} & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

We denote if as $X \sim P(\lambda)$ since $\lambda = np$ is the only parameter.

Our next distribution is finding the number of successes in n trials in the case of a sampling without replacement. Consider an example, among the 100 applicants for a job only 60 are actually qualified. If 4 of the applicants are randomly selected for the next phase of interview and if we are interested in finding the probability that only 2 of the 4 will be qualified, then we proceed as follows.

Entire population (100) can be divided into qualified (60) and non-qualified (40) applicants. The number of ways in which 4 applicants can be selected from the population is $\binom{100}{4}$. Similarly, there are $\binom{60}{2}$ ways of selecting 2 qualified candidates and naturally there are $\binom{40}{2}$ ways of selecting the other 2 candidates.

Hence, the required probability is

$$\frac{\binom{60}{2} \times \binom{40}{2}}{\binom{100}{2}}$$

Infact such problems were already dealt in Chapter 1.

To obtain a formula for the probability of getting x successes in n trials (but the sampling is without replacement) where M of the N (population) elements have a property (in our example, it is 60 qualified applicants). Hence, there are $\binom{M}{x}$ ways of choosing x of the successes and $\binom{N-M}{n-x}$ ways of choosing $n-x$ of the $N-M$ failures. Hence out of $\binom{N}{n}$ ways getting n elements from N, we have x successes and $n-x$ failures. This idea is formulated so that the number of successes in trials is a random variable called the *hypergeometric distribution*.

Definition

A random variable X has a **hypergeometric distribution** if the probability distribution of X is

$$f(x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & \begin{array}{c} x = 0, 1, 2, \cdots, n \text{ and} \\ x \le M \text{ and } n - x \le N - M \end{array} \\ 0 & \text{elsewhere} \end{cases}$$

We denote it as $X \sim hy\,(x; N, M, n)$.

Preceeding discussions can be illustrated by a set of examples we consider now.

4.3 PROBLEMS

■ Example 4.1

10 coins are tossed simultaneously. Find the probability of getting (a) 4 heads, (b) atleast 2 heads.

Let the success event be getting a head with $n = 10$ and $p = p(\text{success}) = \frac{1}{2}$.
Hence $X \sim b\left(10, \frac{1}{2}\right)$

$$\begin{aligned} \therefore \quad f(x) &= \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x} \\ &= \begin{cases} \binom{10}{x}\left(\frac{1}{2}\right)^{10} & x = 0, 1, 2, \cdots 10 \\ 0 & \text{elsewhere} \end{cases} \end{aligned}$$

(a) $p(X = 4) = \binom{10}{4}\left(\dfrac{1}{2}\right)^{10} = 0.2051.$

(b)
$$\begin{aligned} p\,(\text{atleast 2 heads}) &= p(X \geq 2) \\ &= 1 - p(X < 2) \\ &= 1 - p(X = 0, 1) \\ &= 1 - [p(X = 0) + p(X = 1)] \\ &= 1 - \left[\binom{10}{0}\left(\frac{1}{2}\right)^{10} + \binom{10}{1}\left(\frac{1}{2}\right)^{10}\right] \\ &= 1 - \left[\left(\frac{1}{2}\right)^{10} \times 11\right] = 0.9893. \end{aligned}$$ ■

■ Example 4.2

A perfect die is thrown 8 times. Find the probability of getting atmost 3 times 5 or 6 as the outcome.

Let the success event be getting 5 or 6 with $n = 8$ and $p = p(\text{success}) = p(\text{getting 5 or 6 in a single throw}) = \frac{1}{3}$, so that $q = 1 - p = \frac{2}{3}$.

Hence $X \sim b\left(8, \dfrac{1}{3}\right)$ and

$$f(x) = \begin{cases} \binom{8}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{8-x} & x = 0, 1, 2, \cdots, 8 \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} p\,(X \text{ is atmost } 3) &= p(X \leq 3) = p(X = 0) + p(X = 1) + p(X = 2) + p(X = 3) \\ &= \binom{8}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^8 + \binom{8}{1}\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^7 + \binom{8}{2}\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^6 \\ &\quad + \binom{8}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^5 = 0.4682 \end{aligned}$$ ■

■ Example 4.3

A ball is drawn from an urn containing 4 white and 6 black balls. A ball is drawn its colour is noted and it is then replaced. What is the probability that of the first five balls drawn exactly two are white?

Here the success event is getting a white ball

$$\therefore \quad p = \frac{4}{10} = \frac{2}{5} \quad \text{so that} \quad q = \frac{3}{5} \quad \text{and} \quad n = 5 \quad \text{and } X \sim b\left(5, \tfrac{2}{5}\right)$$

$$\therefore \quad \text{Required probability is } p\,(X = 2) = \binom{5}{2}\left(\frac{2}{5}\right)^2\left(\frac{3}{5}\right)^3 = 0.3456$$

■

■ Example 4.4

If the number of accidents occuring on a highway each day is a poisson random variable with parameter $\lambda = 3$, what is the probability that (a) no accidents occur today (b) atmost 2 accidents in a given day.

If X is the number of accidents then we have $X \sim P(3)$. Hence, its probability distribution is

$$f(x) = \begin{cases} e^{-3}\frac{3^x}{x!} & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

(a) $p\,(X = 0) = e^{-3}\left(\dfrac{3^0}{0!}\right) = e^{-3} = 0.0497.$

(b)
$$\begin{aligned} p\,(X \text{ is atmost } 2) &= p\,(x \leq 2) \\ &= p\,(X = 0) + p\,(X = 1) + p\,(X = 2) \\ &= e^{-3}\left(\frac{3^0}{0!}\right) + e^{-3}\left(\frac{3}{1!}\right) + e^{-3}\left(\frac{3^2}{2!}\right) \\ &= 0.42319. \end{aligned}$$

■

■ Example 4.5

Two dice are rolled 100 times. Let X be the number of double sixes. Find the probability that exactly 3 times, we get double six. Using (a) binomial distribution, (b) poisson distribution.

Let X be the number of double sixes

$$\therefore \quad p = \frac{1}{36} \quad q = 1 - \frac{1}{36} = \frac{35}{36} \quad \text{and} \quad n = 100$$

(a) If we assume that $X \sim b\left(100, \frac{1}{36}\right)$, then

$$\begin{aligned} p\,(X = 3) &= \binom{100}{3}\left(\frac{1}{36}\right)^3\left(\frac{35}{36}\right)^{100-3} \\ &= 0.2254. \end{aligned}$$

(b) If X follows poisson distribution with

$$\lambda = np = 100 \times \frac{1}{36} = 2.78$$

$\therefore \quad X \sim P(2.78)$ i.e., its pdf is

$$f(x) = \begin{cases} e^{-2.78}\left[\frac{(2.78)^x}{x!}\right] & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

$$\therefore \quad p(X = 3) = e^{-2.78}\left[\frac{(2.78)^3}{3!}\right] = 0.2221.$$

■

■ Example 4.6

If 2% of the calls received by a switchboard are wrong numbers. Determine the probability that among 150 calls received by the Switchboard atleast 2 are wrong numbers.

If X is the number of wrong numbered calls and since $n = 150$ is large $p = 0.02$ is small we consider that $X \sim P(\lambda)$ where

$$\lambda = np = 150 \times 0.02 = 3$$

Hence,
$$f(x) = \begin{cases} e^{-3}\left(\frac{3^x}{x!}\right) & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \therefore \quad \text{Required probability} &= p(X \geq 2) \\ &= 1 - p(X < 2) \\ &= 1 - [p(X = 0) + p(X = 1)] \\ &= 1 - \left[e^{-3} + e^{-3}(3)\right] \\ &= 0.80085. \end{aligned}$$

***Note*:** In Examples 4.5 and 4.6, we know that X takes respectively from 0 to 100 and 0 to 150. But this numbers of trials ($n = 100$ and 150) is considered to be large we don't include them, while we write the range of X in each case. ■

■ Example 4.7

The numbers of weekly breakdowns of a computer is a random variable having a poisson distribution with $\lambda = 1.5$. Find the probabilities that this computer will function for a given week.

(a) Without a break down.

(b) Atmost 3 break down.

(c) Atleast 1 break down.

Let X be the number of break down of the computer. Given $X \sim P(\lambda)$, $\lambda = 1.5$. So that its pdf is

$$f(x) = \begin{cases} e^{-1.5}\left[\frac{(1.5)^x}{x!}\right] & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

(a) $p(X = 0) = e^{-1.5}\left[\dfrac{(1.5)^0}{0!}\right] = 0.2231.$

(b)
$$\begin{aligned} p(x \le 3) &= e^{-1.5}\left[\frac{(1.5)^0}{0!}\right] + e^{-1.5}\left[\frac{(1.5)^1}{1!}\right] + e^{-1.5}\left[\frac{(1.5)^2}{2!}\right] + e^{-1.5}\left[\frac{(1.5)^3}{3!}\right] \\ &= e^{-1.5}[1 + 1.5 + 1.125 + 0.5625] \\ &= 0.9344. \end{aligned}$$

(c)
$$\begin{aligned} p(X \ge 1) &= 1 - p(X = 0) \\ &= 1 - 0.2231 \\ &= 0.7769. \end{aligned}$$

■

■ Example 4.8

If a balanced dia is thrown, find the probability that a ‘6’ first appears on the 7th trial.

If the success event is assume to be getting a ‘6’, then at 7th trial we have the first success. Also $p = \frac{1}{6}$.

Hence if X if the number of trials then $X \sim Ge\left(\frac{1}{6}\right)$.

$$\therefore \quad f(x) = \begin{cases} \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \text{Here} \quad x + 1 &= 7 \\ \therefore \quad x &= 6 \\ \therefore \quad \text{Required probability} &= \frac{1}{6}\left(\frac{5}{6}\right)^6 \\ &= 0.0558. \end{aligned}$$

***<u>Note</u>*:** In Example 4.8 if we wish to find the probability that a ‘6’ appears 3rd time on the 7th trial then the random variable X follows a negative binomial distribution with $x = 4$, $r = 3$ and $p = \frac{1}{6}$ Hence the required probability

$$\begin{aligned} &= \binom{6}{2}\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)^4 \\ &= 0.0335. \end{aligned}$$

■

■ Example 4.9

If a boy is throwing at a target. What is the probability that his 12th throw is his 7th hit if the probability of hitting that target is 0.4?

If the success event is the hitting the target then given that $p = 0.4$ and he is registering his 7^{th} hit $(= r)$ at his $12^{\text{th}}(= x + r)$ hit. Hence X follows negative binomial distribution with $p = 0.4$, $r = 7$. So that

$$f(x) = \begin{cases} \begin{pmatrix} x+7-1 \\ 7-1 \end{pmatrix} (0.4)^7 (0.6)^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

However, $x + r = 12$ (or) $x + 7 = 12$ implies that $x = 5$.

$$\begin{aligned} \therefore \quad \text{Required probability is } p\,(X = 5) &= \begin{pmatrix} 11 \\ 6 \end{pmatrix} (0.4)^7\,(0.6)^5 \\ &= 0.0589. \end{aligned}$$

■

■ Example 4.10

If the probability is 0.75 that a sales person will convince his customer to sell his product, find the probabilities that (a) the eigth customer will be the fifth customer convinced by the sales person, (b) the fifteenth customer will be the tenth customer convinced by the sales person.

If the success event is the customer convinced by the sales person then $p = 0.75$.

In both cases, we observe that the random variable X follows a negative binomial distribution with $p = 0.75$, so that $q = 1 - p = 0.25$.

(a) $X \sim nb\,(x; 5, 0.75)$

$$\therefore \quad f(x) = \begin{cases} \begin{pmatrix} x+5-1 \\ 5-1 \end{pmatrix} (0.75)^5 (0.25)^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Also $x + r = x + 5 = 8$ implies that $x = 3$. Hence, the required probability is

$$\begin{aligned} p\,(X = 3) &= \begin{pmatrix} 7 \\ 4 \end{pmatrix} (0.75)^5\,(0.25)^3 \\ &= 0.1298. \end{aligned}$$

(b) $X \sim nb(x; 10, 0.75)$

$$f(x) = \begin{cases} \begin{pmatrix} x+10-1 \\ 10-1 \end{pmatrix} (0.75)^{10} (0.25)^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Here $x + r = 15$ (or) $x + 10 = 15$, so that $x = 5$

$$\begin{aligned} \therefore \quad p\,(X = 5) &= \begin{pmatrix} 14 \\ 9 \end{pmatrix} (0.75)^{10}\,(0.25)^5 \\ &= 0.1101. \end{aligned}$$

■

■ Example 4.11

Among the 200 employees of a company, 120 are post graduates. If 8 of the employees are chosen by a lot, find the probability that (a) 4 of the eight will be post graduates, (b) majority are non-post graduate employees.

If $N = 200$ is the total number of employees and $M = 120$ are post graduates, then $N-M = 80$ are non-Post graduates. Sample size is $n = 8$. Let X be the number of post graduates in the sample.

$\therefore \quad X \sim hy(x; 200, 120, 8)$. So that is pdf is

$$f(x) = \begin{cases} \dfrac{\begin{pmatrix} 120 \\ x \end{pmatrix}\begin{pmatrix} 80 \\ 8-x \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} & x = 0, 1, 2, 3, \cdots, 8 \\ 0 & \text{elsewhere} \end{cases}$$

(a) $p(x = 4) = \dfrac{\begin{pmatrix} 120 \\ 4 \end{pmatrix}\begin{pmatrix} 80 \\ 4 \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} = 0.2358.$

(b) p(Majority are non-post graduate employees)

$$\begin{aligned} &= p\,(X = 0) + p\,(X = 1) + p\,(X = 2) + p\,(X = 3) \\ &= \frac{\begin{pmatrix} 120 \\ 0 \end{pmatrix}\begin{pmatrix} 80 \\ 8 \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} + \frac{\begin{pmatrix} 120 \\ 1 \end{pmatrix}\begin{pmatrix} 80 \\ 7 \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} + \frac{\begin{pmatrix} 120 \\ 2 \end{pmatrix}\begin{pmatrix} 80 \\ 6 \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} + \frac{\begin{pmatrix} 120 \\ 3 \end{pmatrix}\begin{pmatrix} 80 \\ 5 \end{pmatrix}}{\begin{pmatrix} 200 \\ 8 \end{pmatrix}} = 0.0464 \end{aligned}$$

***Note*:** Since n is very small compared to N in Examples 4.11, we can use binomial distribution to approximate hypergeometric distribution.

That is, if we assume that selecting a post graduate employee is 'success' event then $p = \frac{M}{N} = \frac{120}{200} = 0.6$, $q = 1 - p = 0.4$ and $n = 8$. We shall compute the probablity in the first part of Example 4.11.

$\therefore \quad X \sim b(8, 0.6)$ and its pdf is

$$f(x) = \begin{cases} \begin{pmatrix} 8 \\ x \end{pmatrix} (0.6)^x (0.4)^{8-x} & x = 0, 1, 2, \cdots, 8 \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \text{Required probability} &= p\,(X = 4) = \begin{pmatrix} 8 \\ 4 \end{pmatrix} (0.6)^4 (0.4)^4 \\ &= 0.2322. \end{aligned}$$

Such approximation can be used in general if n is relatively small compared to N and a preferable thumb rule is if $n < \frac{N}{20}$. ■

■ Example 4.12

If an income tax official selects 4 returns from 15 returns of which 7 contain illegitimate deductions, find the probability that he will find 3 income-tax returns with illegitimate deductions.

Here $N = 15$ $M =$ number of returns with illegitimate deductions $= 7$ then $N - M = 8$. Also $n = 4$ and X can be the number of returns with illegitimate deductions in the sample.

$\therefore \quad X \sim hy\,(x;\,15, 7, 4)$ so that its pdf is

$$f(x) = \begin{cases} \dfrac{\binom{7}{x}\binom{8}{4-x}}{\binom{15}{4}} & x = 0, 1, 2, \cdots, 7 \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{Required probability is} = p\,(X = 3) = \frac{\binom{7}{3}\binom{8}{1}}{\binom{15}{4}}$$

$$= 0.2051.$$

■

■ Example 4.13

A throws 2 dice until he gets 6 and B throws independently 2 other dice until he gets 7. Find the probability that B will require more throws than A.

Let X be the number of trials required by A to get his first success. Then $X \sim Ge\,(p_1)$ where $p_1 = p(\text{getting six when two dice are thrown}) = \frac{5}{36}$.

Hence $q_1 = 1 - p_1 = \frac{31}{36}$ and the pmf of X is

$$f_X(x) = \begin{cases} \left(\frac{5}{36}\right)\left(\frac{31}{36}\right)^x & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Similarly if Y is the number of trials for B to get his first success, then $y \sim Ge\,(p_2)$ where $p_2 = p(\text{getting seven when two dice are thrown})$

$$= \frac{6}{36} \text{ or } \frac{1}{6} \quad \text{so that} \quad q_2 = 1 - \frac{1}{6} = \frac{5}{6}$$

$$\text{pmf of } Y \text{ is } f_Y(y) = \begin{cases} \frac{1}{6}\left(\frac{5}{6}\right)^y & y = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Required probability is $p(B \text{ requires more throws than } A)$.

$$= p\,(X = 0)\,p\,(y = 1, 2, 3, \cdots) + p\,(X = 1) \cdot p\,(y = 2, 3, 4, \cdots) + \cdots$$

($\because$ X and Y are independent random variables)

$$
\begin{aligned}
&= p\,(X=0)\,p\,(y>0) + p\,(X=1)\,p\,(y>1) + p\,(X=2)\,p\,(y>2) + \cdots \\
&= \left(\frac{5}{36}\right)\sum_{y=1}^{\infty}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^4 + \left(\frac{5}{36}\right)\left(\frac{31}{36}\right)\sum_{y=2}^{\infty}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^y + \left(\frac{5}{36}\right)\left(\frac{31}{36}\right)^2\sum_{y=3}^{\infty}\left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^4 + \cdots \\
&= \left(\frac{5}{36}\right)\left(\frac{1}{6}\right)\left[\sum_{y=1}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)\sum_{y=2}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)^2\sum_{y=3}^{\infty}\left(\frac{5}{6}\right)^4 + \cdots\right] \\
&= \left(\frac{5}{216}\right)\left[\left(\frac{5}{6}\right)\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)\left(\frac{5}{6}\right)^2\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)^2\left(\frac{5}{6}\right)^3\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \cdots\right] \\
&= \left(\frac{5}{216}\right)\left(\frac{5}{6}\right)\left[\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)\left(\frac{5}{6}\right)\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \left(\frac{31}{36}\right)^2\left(\frac{5}{6}\right)^2\sum_{y=0}^{\infty}\left(\frac{5}{6}\right)^y + \cdots\right] \\
&= \frac{25}{1296}\left[\left(\frac{1}{1-5/6}\right) + \frac{155}{216}\left(\frac{1}{1-5/6}\right) + \left(\frac{155}{216}\right)^2\left(\frac{1}{1-5/6}\right) + \cdots\right] \\
&= \left(\frac{25}{1296}\right)\left(\frac{1}{1-5/6}\right)\left[1 + \frac{155}{216} + \left(\frac{155}{216}\right)^2 + \cdots\right] \\
&= \left(\frac{25}{1296}\right)\left(\frac{1}{1/6}\right)\left(\frac{1}{1-\frac{155}{216}}\right) = \left(\frac{25}{216}\right)\left(\frac{216}{61}\right) = \frac{25}{61}
\end{aligned}
$$

■

4.4 CONTINUOUS DISTRIBUTIONS

In this section several parametric families of probability density functions are presented. We shall start with a very simple distribution for a continuous random variable. A random variable X which has a constant probability over an internal (a, b) where $a < b$ and a and b are real numbers. One can easily verify that the constant is $\left(\frac{1}{b-a}\right)$, hence we have the following definition.

Definition

A random variable X has a **uniform distribution** if the probability density function is

$$
f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{elsewhere} \end{cases}
$$

We denote X as $X \sim u(a, b)$ and a and b are the parameters.

Even though we define X as being uniformly distributed over the open interval (a, b) we could define it over the closed internal $[a, b]$ or over either of $(a, b]$ or $[a, b)$. One of the applications of uniform distribution we shall find in the future discussion in transformation of variables.

Uniform distribution is also called the *rectangular* distribution, since the shape of the density is rectangular.

Our next distribution namely normal distribution which is one of the more widely used distributions in applications of statistical methods. Variables such as length of steel rods, the score of a cricket player, and life of an electric lamp are often assumed to be random variables having normal definition. Let us define more formally now and have a detailed discussion of some of the important aspects of normal distribution.

Definition

A random variable X is said to have a normal distribution if X has a porbability density function

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

We denote it by $X \sim n\left(\mu, \sigma^2\right)$ where μ and σ^2 are the parameters.

***Note*:** If the parameters are assumed as $\mu = 0$ and $\sigma = 1$, we have a special case of X which is referred to as the standard normal distribution whose pdf is

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad -\infty < x < \infty$$

Details of the special case will be studied in the subsequent chapter where as its utility in applied calculations will be discussed in this chapter itself especially in the calculation of probabilities involving $X \sim n\left(\mu, \sigma^2\right)$ and where we avoid (or impossible) a direct integration.

Other family of distribution that plays important roles in statistics is gamma distribution from which we get two more distributions as special cases, namely exponential distribution and chi-square distribution. The exponential distribution has been widely used as a model for lifetimes of various things such as *reliability*. Where as chi-square (particularly central chi-square) distribution finds a wide application and utility in testing of hypothesis that is in sampling theory, which is not listed in our discussion. But a formal definition of gamma distribution is presented here.

Definition

A random variable X has a **gamma distribution** if its probability density function is

$$f(x) = \begin{cases} \frac{1}{\sqrt{\alpha}\beta^\alpha} x^{\alpha-1} e^{-x/\beta} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

We denote it by $X \sim G(X; \alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$ are the parameters. When α is not a positive integer, the value of $\lfloor\overline{\alpha}$ will have to be referred in a special table.

In gamma distribution if $\alpha = 1$ and $\beta = 1/\lambda$, we get an exponential distribution.

Definition

A random X is said to have an **exponential distribution** if X has a probability density function.

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

We denote if by $X \sim Expo(\lambda)$ where λ is the only parameter.

Another special case of gamma distribution namely **chi-square distribution** is obtained by taking $\alpha = r/2$ and $\beta = 2$.

Definition

A random variable X is said to have a chi-square distribution, if X has a probability density function

$$f(x) = \begin{cases} \frac{1}{2^{r/2}\Gamma\left(\frac{r}{2}\right)} x^{\frac{r}{2}-1} e^{-x/2} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

We denote it by $X \sim \chi^2(r)$ and r is the parameter which is referred to as the number of degrees of freedom.

Gamma and exponential distributions can be thought of as a continuous waiting time random variable so that the negative binomial and geometric distributions are the discrete analogs of these two distributions.

Now let us define another useful distribution which has a wide application in *reliability* theory (lifetime of an item). We call it as *weibull* distribution and define it as follows.

Definition

A random variable X has a **weibull distribution** if the probability density of X is

$$f(x) = \begin{cases} (\alpha\beta)x^{\beta-1}e^{-\alpha x^{\beta}} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

We denote it by $X \sim We(x; \alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$ are the parameters. Observe that with $\beta = 1$, we get exponential distribution with parameter α.

Beta distribution is our next variable whose application can be found in Bayesian inference and which is quite flexible as beta distribution can takes great variety of different shapes so that the it can be used to model an experiment for which one of these shapes is appropriate.

Definition

A random variable X is said to have a **beta distribution** if X has a probability density function

$$f(x) = \begin{cases} \frac{1}{\beta(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

We denote it by $X \sim Be(x; \alpha, \beta)$ where $\alpha > 0$ and $\beta > 0$ are the parameters of a beta distribution.

If $\alpha = 1$ and $\beta = 1$, we can observe that beta distribution reduces to the uniform distribution $f(x) = 1$ in (0, 1). This definition can also be presented using Gamma functions as

$$f(x) = \begin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Let us consider some fundamental examples related to the above listed special continuous distributions.

4.5 PROBLEMS

■ Example 4.14

Trains arrive at a specified station at 15 minute intervals starting from 9 a.m. If a passenger arrives at the station in a random time between 9 and 9.30 a.m. Find the probability that he has to wait.

(a) less than 5 minutes,

(b) atleast 12 minutes for the train.

Let X be the waiting time of the passenger so that X takes any value between 0 and 30 (since he enters the station any time between 9 and 9.30 a.m.).

Hence, $X \sim u(0, 30)$ and its pdf is

$$f(x) = \begin{cases} \frac{1}{30} & 0 < x < 30 \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \text{(a)} \quad p(X < 5) &= p(\text{he arrives either between 9.10 and 9.15 a.m. or 9.25 and 9.30 a.m.}) \\ &= p(10 < X < 15) + p(25 < X < 30) \\ &= \int_{10}^{15} \left(\frac{1}{30}\right) dx + \int_{25}^{30} \left(\frac{1}{30}\right) dx = \frac{1}{3} \\ \text{(b)} \quad p(X > 12) &= p(0 < X < 3) + p(15 < X < 18) \\ &= \int_{0}^{3} \left(\frac{1}{30}\right) dx + \int_{15}^{18} \left(\frac{1}{30}\right) dx = \frac{1}{5} \end{aligned}$$

■

■ Example 4.15

A random variable $X \sim u(-2, 2)$ compute (a) $p(X < 1)$, (b) $p(|X| < 1/2)$, (c) $p(|X - 1| < 1)$ and (d) Find K such that $p(X > K) = \frac{1}{4}$.

Given $X \sim u(-2, 3)$. Its pdf is

$$f(x) = \begin{cases} \frac{1}{4} & -2 < x < 2 \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{(a)} \quad p(X < 1) = \int_{-2}^{1} \left(\frac{1}{4}\right) dx = \frac{1}{4}(x)'_{-2} = \frac{3}{4}$$

$$\begin{aligned} \text{(b)} \quad p(|X| < 1/2) &= p\left(-\frac{1}{2} < x < \frac{1}{2}\right) = \int_{-1/2}^{1/2} \left(\frac{1}{4}\right) dx \\ &= \frac{1}{4}(x)_{-1/2}^{1/2} = \frac{1}{4} \end{aligned}$$

(c) $$p(|X-1|<1) = p(-1<X-1<1) = p(0<X<2)$$
$$= \int_0^2 f(x)\,dx = \int_0^2 \left(\frac{1}{4}\right) dx = \frac{1}{2}$$

(d) given
$$p(X>K) = \frac{1}{4}$$
$$\Rightarrow \int_K^2 f(x)\,dx = \frac{1}{4} \quad \Rightarrow \int_K^2 \left(\frac{1}{4}\right) dx = \frac{1}{4}$$
$$\Rightarrow \frac{1}{4}(x)_K^2 = \frac{1}{4}$$
$$\Rightarrow \frac{1}{4}(2-K) = \frac{1}{4}; \qquad K = 1$$

■

■ Example 4.16

While finding the thickness of a certain wire, the error made is a random variable which follows uniform density from -0.01 and 0.01. Find the probabilities that such an error will (a) be between -0.002 and 0.003 and, (b) exceed 0.005 in absolute value.

Given $X \sim u(-0.01, 0.01)$, hence its pdf is

$$f(x) = \begin{cases} \frac{1}{0.02} & -0.01 < X < 0.01 \\ 0 & \text{elsewhere} \end{cases}$$

(a) $$p(-0.002<X<0.003) = \int_{-0.002}^{0.003} f(x)\,dx = \int_{-0.002}^{0.003} \left(\frac{1}{0.02}\right) dx$$
$$= \left(\frac{1}{0.02}\right)(x)_{-0.002}^{0.003} = 0.25.$$

(b) $$p(|X|>0.005) = 1 - p(|X| \le 0.005) = 1 - p(-0.005 \le X \le 0.005)$$
$$= 1 - \int_{-0.005}^{0.005} f(x)\,dx = 1 - \int_{-0.005}^{0.005} \left(\frac{1}{0.02}\right) dx$$
$$= 1 - (x)_{-0.005}^{0.005} = 0.99.$$

■

■ Example 4.17

On a certain route, buses run every 45 minutes between mid-night and six in the morning. What is the probability that a man entering the bus station at a random time during this period will have to wait atleast K minutes with $K = 5, 10, 15, 20, 30$?

Let X be the waiting time of the person to get the bus. Since buses run every 45 minutes $X \sim u(0, 45)$.

$\therefore$ Its pdf is

$$f(x) = \begin{cases} \frac{1}{45} & 0 < x < 45 \\ 0 & \text{elsewhere} \end{cases}$$

$$p(X \geq K) = \int_K^\infty f(x)\,dx = \int_K^{45} \left(\frac{1}{45}\right) dx$$

$$= \left(\frac{1}{45}\right) (x)_K^{45} = \frac{45 - K}{45}$$

Substituting $K = 5, 10, 15, 20, 30$ we get the respective probabilities as,

(a) $p(X \geq 5) = \dfrac{45 - 5}{45} = 0.8889.$

(b) $p(X \geq 10) = \dfrac{45 - 10}{45} = 0.7778.$

(c) $p(X \geq 15) = \dfrac{45 - 15}{45} = 0.6667.$

(d) $p(X \geq 20) = \dfrac{45 - 20}{45} = 0.5556.$

(e) $p(X \geq 30) = \dfrac{45 - 30}{45} = 0.3333.$ ■

■ Example 4.18

In a certain city, the daily consumption of electric power, in Millions of Killowatt hours (MKh) may be regarded as a random variable having a gamma distribution with $\alpha = 3$ and $\beta = 2$. If the power plant has a daily capacity of 12 MKh, what is the probability that this power supply will be inadequate on any given day?

Let X be the daily consumption of electric power.
Given that $X \sim G(x; 3, 2)$
$\therefore$ its pdf is

$$f(x) = \begin{cases} \left(\frac{1}{\lceil 3 \; 2^3}\right) x^{3-1} e^{-x/2} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

(or)

$$f(x) = \begin{cases} \left(\frac{1}{48}\right) x^2 e^{-x/2} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$
\begin{aligned}
p(\text{power supply is inadequate}) &= p(X > 12) \\
&= \int_{12}^{\infty} f(x)dx = \int_{12}^{\infty} \left(\frac{1}{48}\right) x^2 e^{-x/2}\, dx \\
&= \frac{1}{48}\left[x^2\left(\frac{e^{-x/2}}{-1/2}\right) - 2x\left(\frac{e^{-x/2}}{1/4}\right) + 2\left(\frac{e^{-x/2}}{-1/8}\right)\right]_{12}^{\infty} \\
&= \frac{1}{48}\left[0 - e^{-6}\left(-288 - 96 - 16\right)\right] \\
&= 0.0208.
\end{aligned}
$$

■

■ Example 4.19

The lifetime of an electric bulb is a random variable having an exponential distribution with $\lambda = 0.01$. Find the probabilities that such a bulb will last (a) atleast 120 days, (b) atmost 200 days and (c) anywhere between 75 and 100 days.

Given $X \sim Expo(\lambda)$ with $\lambda = 150$ and hence its pdf is

$$
f(x) = \begin{cases} (0.01)e^{-0.01x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}
$$

$$
\begin{aligned}
\text{(a)}\quad p\left(X \geq 120\right) &= \int_{120}^{\infty} f(x)dx = \int_{120}^{\infty} (0.01)e^{-0.01x}\, dx \\
&= 0.01\left(\frac{e^{-0.01x}}{-0.01}\right)_{120}^{\infty} \\
&= -\left(0 - e^{-1.2}\right) = 0.3012.
\end{aligned}
$$

$$
\begin{aligned}
\text{(b)}\quad p\left(X \leq 200\right) &= \int_{-\infty}^{200} f(x)dx = \int_{0}^{200} (0.01)e^{-0.01x}\, dx \\
&= 0.01\left(\frac{e^{-0.01x}}{-0.01}\right)_{0}^{200} \\
&= \left(1 - e^{-2}\right) = 0.8647.
\end{aligned}
$$

$$
\begin{aligned}
\text{(c)}\quad p\left(75 < X < 100\right) &= \int_{75}^{100} (0.01)e^{-0.01x}\, dx \\
&= -\left(e^{-0.01x}\right)_{75}^{100} = e^{-0.75} - e^{-1} \\
&= 0.1045.
\end{aligned}
$$

■

■ Example 4.20

If the annual proportion of a component that fail in a certain brand of television set may be looked upon as a random variable having a beta distribution with $\alpha = 2$ and $\beta = 4$ find the probability that atleast 25% of all that component will fail in the television sets of that brand.

If X is the annual proportion of a component that fail, then $X \sim Be\,(x; 2, 4)$

$\therefore$ Its pdf is

$$f(x) = \begin{cases} \frac{1}{\beta(2,4)} x^{2-1}\,(1-x)^{4-1} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{(or)} \qquad f(x) = \begin{cases} \frac{\lceil 6}{\lceil 2\ \lceil 4} x(1-x)^3 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

$$= \begin{cases} 20x(1-x)^3 & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Required probability is $p(X \geq 0.25)$

$$= \int_{0.25}^{\infty} f(x)\,dx = 20 \int_{0.25}^{1} x\,(1-x)^3\,dx$$

$$= 20 \int_{0.25}^{1} x\left(1 - 3x + 3x^2 - x^3\right)dx$$

$$= 20\left(\frac{x^2}{2} - x^3 + \frac{3}{4}x^4 - \frac{x^5}{5}\right)_{0.25}^{1}$$

$$= 20\left[\left(\frac{1}{2} - \frac{0.25^2}{2}\right) - \left(1 - 0.25^3\right) + \frac{3}{4}\left(1 - 0.25^4\right) - \frac{1}{5}\left(1 - 0.25^5\right)\right]$$

$$= 0.6328.$$

■

■ Example 4.21

If the life (in hours) of a semiconductor is a random variable having a weibull distributions with $\alpha = 0.025$ and $\beta = 0.5$. Find the probability that such a semiconductor will still be in operating condition after 4,000 hours?

Let X be the life (in hours) of the semiconductor which follows weibull distribution.

Given that $X \sim We(x; 0.025, 0.5)$. Hence, its pdf is

$$f(x) = \begin{cases} (0.025)\,(0.5)x^{(0.5-1)}e^{-0.025(x)^{0.5}} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

(or) $$f(x) = \begin{cases} (0.0125)x^{-0.5}e^{-0.025(x)^{0.5}} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$\therefore \text{ Required probability is } p\,(X > 4000) = \int_{4000}^{\infty} f(x)dx = \int_{4000}^{\infty} (0.0125)x^{-0.5}e^{-0.025(x)^{0.5}}dx$$

Let $0.025x^{(0.5)} = Z$ so that Z takes values from 1.58 to ∞ and the integral becomes

$$\begin{aligned} p\,(X > 4000) &= p\,(Z > 1.58) = \int_{1.58}^{\infty} e^{-Z}\,dz \\ &= \left(-e^{-Z}\right)_{1.58}^{\infty} \\ &= e^{-1.58} \\ &= 0.206. \end{aligned}$$

■

■ Example 4.22

The time (in hours) required to repair a machine is exponentially distributed with parameter $\lambda = \frac{1}{4}$. (a) What is the probability that the repair time exceeds 1 hour? (b) If the repair time has already lasted for 8 hours what is the probability that it will last for atleast two more hours?

Let X be the time to repair the machine.
Given $X \sim Expo(\lambda)$ with $\lambda = \frac{1}{4}$, the density function of X is given by

$$f(x) = \begin{cases} \left(\frac{1}{4}\right) x^{-x/4} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$\text{(a)}\quad \begin{aligned} p\,(X > 1) &= \int_{1}^{\infty} f(x)\,dx = \int_{1}^{\infty} \left(\frac{1}{4}\right) e^{-x/4}dx \\ &= e^{-1/4} \\ &= 0.7788. \end{aligned}$$

$$
\begin{aligned}
\text{(b)} \quad p\,(X > 10/X > 8) &= \frac{p\,(X > 10 \cap X > 8)}{p\,(x > 8)} \\
&= \frac{p\,(X > 10)}{p\,(X > 8)} \\
&= \frac{e^{-5/2}}{e^{-2}} = e^{-1/2} \\
&= 0.6065.
\end{aligned}
$$

Remark

In Example 4.22 (b) shows a property of exponential distribution called *memory less property* which we shall discuss later. However, here we used the definition of conditional events and their probabilities laws.

■ Example 4.23

A component has the failure distribution

$$
f(x) = \begin{cases} (0.4x)e^{-0.2x^2} & x \geq 0 \\ 0 & \text{elsewhere} \end{cases}
$$

where X is the time to failure measured in years. Determine the probability that the component failing with in the first month of its operation.

Let X be the time to failure in years. Given pdf of X is

$$
\begin{aligned}
f(x) &= \begin{cases} 0.4xe^{-0.2x^2} & x \geq 0 \\ 0 & \text{elsewhere} \end{cases} \\
&= \begin{cases} (0.2)2x^{2-1}\, e^{-0.2x^2} & x > 0 \\ 0 & \text{elsewhere} \end{cases}
\end{aligned}
$$

Hence $X \sim We(x; 0.2, 2)$.
The required probability is

$$
\begin{aligned}
p(X < 1 \text{ month}) &= p\left(X < \frac{1}{12}\right) \\
&= \int_{-\infty}^{1/12} f(x)\, dx = \int_{0}^{1/12} f(x)\, dx \\
&= \int_{-\infty}^{1/12} (0.4x)\, e^{-0.2x^2} dx
\end{aligned}
$$

Let $0.2x^2 = Z$ so that $0.4x\ dx = dz$ and Z varies from 0 to 0.0014.

$$\begin{aligned} \therefore \qquad p(X < 1 \text{ month}) &= \int_0^{0.0014} e^{-z}\, dz = \left(-e^{-z}\right)_0^{0.0014} \\ &= 1 - e^{-0.0014} = 0.0014. \end{aligned}$$

Remark

If the random variable T is the time to failure of an item and $f(t)$ and $F(t)$ are the pdf and cdf of T respectively, then its failure rate is defined as the probability density of failure at time t given that failure does not occur prior to time t. The conditional probability of failure per unit time is given as $\frac{f(t)}{1-F(t)}$.

That is, the conditional probability of failure in the interval $(t, t+\Delta t)$, given that the item has survived upto time t, is

$$\begin{aligned} p(t \le T \le t + \Delta t/T \ge t) &= \frac{p(t \le T \le t + \Delta t)}{p(T \ge t)} \\ &= \frac{p(T \le t + \Delta t) - p(T \ge t)}{p(t \ge t)} \\ &= \frac{F(t+\Delta t) - F(t)}{1 - p(T < t)} \\ &= \frac{\Delta t}{1 - F(t)} \frac{F(t+\Delta t) - F(t)}{\Delta t} \\ &= \frac{p\,(t \le T \le t + \Delta t/T \ge t)}{\Delta t} = \frac{f(t)}{1 - F(t)} \end{aligned}$$

as $\Delta t \to 0$ and since $F'(t) = f(t)$

Hence, the failure rate is $\frac{f(t)}{1-F(t)}$. In particular, if T has an exponential distribution, then $f(t) = \lambda e^{-\lambda t}$ and $F(t) = 1 - e^{-\lambda t}$ so that the failure rate is $\frac{\lambda e^{-\lambda t}}{1-(1-e^{-\lambda t})} = \lambda$ a constant. As an another example, if $T \sim$ we $(t; \alpha, \beta)$ then $f(t) = \alpha\beta t^{-1}\, e^{-\alpha t^{\beta}}$ and $F(t) = 1 - e^{-\alpha t^{\beta}}$ and hence the failure rate is $\frac{\alpha\beta t^{\beta-1} e^{-\alpha t^{\beta}}}{1-(1-e^{-\alpha t^{\beta}})} = \alpha\beta t^{\beta-1}$.

Indeed converse of the above results are true but we omit the details of the proof here. These results play important role in reliability theory.

Further such failure rate for a discrete case, that is, on the x^{th} trial, then its failure rate at the x^{th} trial is the probability that it will fail on the x^{th} trail given that it has not failed on the first $x-1$ trials which is given by $\frac{f(x)}{1-F(x-1)}$. Analogus to exponential distribution if X is a geometric random variable, its failure rate is constant.

4.6 MORE ON NORMAL DISTRIBUTION

Each of the special distributions considered so far has been illustrated with some examples. Indeed for the normal distribution such an attempt will be made in this section. This idea of discussion has two reasons. the normal distribution is one of the more widely used distributions in applications of statistical methods. Secondly, the calculation of probability of an event described in terms of a normal random variable X, will not make use of direct integration techniques. In fact we start a very useful theorem to justify the second reason and utilise it in the subsequent examples. Recall the following definition.

Definition

A random variable X is said to have a normal distribution if it has a pdf of the form.

$$f(x) = \frac{1}{e\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

In our earlier section we observed another pdf derived from the above one is the pdf of standard normal variable usually denoted by 'Z'. Following Theorem states the method of obtaining Z from X. We do the proof of this Theorem in Chapter 6. But the utility of this Theorem 4.1 can be enjoyed in this section itself.

Theorem 4.1

If the random variable X is $n(\mu, \sigma^2)$ where $\sigma^2 > 0$, then the random variable $Z = \frac{X-\mu}{\sigma}$ is $n(0, 1)$.

Using the properties of expectations we can easily observe that

$$\begin{aligned} E(Z) &= E\left(\frac{X-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma}[E(X) - \mu] = \frac{1}{\sigma}(\mu - \mu) = 0 \\ V(Z) &= V\left(\frac{X-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma^2}V(X) = \frac{1}{\sigma^2} \cdot \sigma^2 = 1 \end{aligned}$$

The above Theorem 4.1 simplifies calculations of probabilities concerning normal variables. Let us start with an example.

■ Example 4.24

If X is $n(75, 100)$ find $p(X < 60)$ first, let us try a direct procedure using $X \sim n(75, 100)$. Here $\mu = 75, \sigma^2 = 100$ and $\sigma = 10$.

$\therefore$ pdf of X is

$$f(x) = \frac{1}{10\sqrt{2\pi}} e^{-\frac{(x-75)^2}{200}} \quad -\infty < x < \infty$$

Hence,
$$p(X < 60) = \int_{-\infty}^{60} \frac{1}{10\sqrt{2\pi}} e^{-\frac{(x-75)^2}{200}} dx$$

This integral cannot be evaluated directly by the basic integration techniques. Instead, if we use Theorem 4.1, the given probability concerning $X \sim n(75, 100)$ can be expressed in terms of probabilities concerning $Z \sim n(0, 1)$ we shall solve examples after observing a set of results which effectively changes X to Z.

Result 1: Let $X \sim n(\mu, \sigma^2)$ a and b are two constants such that $a < b$. Then

$$\begin{aligned}
p(a < X < b) &= p(X < b) - p(X < a) \\
&= p\left(\frac{X-\mu}{\sigma} < \frac{b-\mu}{\sigma}\right) - p\left(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right) \\
&= p\left(Z < Z_2\right) - p\left(Z < Z_1\right) \\
&\quad \text{where} \quad Z_1 = \frac{a-\mu}{\sigma}, \ \text{and} \ Z_2 = \frac{b-\mu}{\sigma} \\
&= \int_{-\infty}^{Z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}\, dz - \int_{-\infty}^{Z_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}\, dz.
\end{aligned}$$

Result 2: Let $X \sim n\left(\mu, \sigma^2\right)$. If a is any number, then

$$\begin{aligned}
p(X > a) &= 1 - p(X < a) \\
&= 1 - p\left(\frac{X-\mu}{\sigma} < \frac{a-\mu}{\sigma}\right) \\
&= 1 - p\left(Z < Z_1\right) \\
&= 1 - \int_{-\infty}^{Z_1} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}\, dz
\end{aligned}$$

Hence probabilities concerning X can be expressed in terms of Z. However, an integral of the form $\int_{-\infty}^{K} \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}} dz$ cannot be evaluated. But, tables of approximate value of this integral for various values of K have been prepared and are reproduced (may not be fully) in Appendix. In this way, indeed we avoid direct integration.

■ Example 4.24 (Continued)

Given $X \sim n(75, 100)$ so that $\mu = 75$ and $\sigma = 10$.

$$\begin{aligned}
\text{Let} \qquad Z &= \frac{X-\mu}{\sigma} = \frac{X-75}{10} \quad \text{and} \quad Z \sim n(0,1) \\
\text{Now} \qquad p(X < 60) &= p\left(\frac{X-75}{10} < \frac{60-75}{10}\right) \\
&= p(Z < -1.5) = 0.0668
\end{aligned}$$

We now state some important properties of the normal distribution and normal probability curve.

(a) The curve is bell shoped and symmetrical about the line $X = \mu$ that is, $p(-\infty < X < \mu) = p(\mu < X < \infty)$. However $p(-\infty < X < \infty) = 1$.

$$\begin{aligned}
\Rightarrow \quad & p(-\infty < X < \mu) + p(\mu < X < \infty) = 1 \\
\Rightarrow \quad & 2p(-\infty < X < \mu) = 1 \\
\Rightarrow \quad & p(-\infty < X < \mu) = \frac{1}{2} = p(\mu < X < \infty) \text{ and hence} \\
& p(-\infty < z < 0) = p(0 < z < \infty) = \frac{1}{2}
\end{aligned}$$

(b) $f(x)$ has maximum probability at the point $x = \mu$ and given by

$$[f(x)]_{\max} = \frac{1}{\sigma\sqrt{2\pi}}$$

(c) X-axis is an asymptote to the curve of $f(x)$.

(d) Coefficient of skewness of X is zero and the coefficient of kurtosis of X is three.

(e) Area property: The area under the normal probability curve between $(\mu - \sigma, \mu + \sigma)$ is 0.6826.

$$\begin{aligned}
\text{Since,} \quad p(\mu - \sigma < X < \mu + \sigma) &= p(-1 < Z < 1) \\
&= 2p(0 < Z < 1) \qquad \text{(using symmetric property)} \\
&= 0.6826
\end{aligned}$$

In a similar way, we have

$$\begin{aligned}
p(\mu - 2\sigma < X < \mu + 2\sigma) &= 0.9544 \quad \text{and} \\
p(\mu - 3\sigma < X < \mu + 3\sigma) &= 0.9973
\end{aligned}$$

Using the symmetric property the table of standard normal distribution can be modified as the value of $\int_0^K \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$. For convenience we present these both forms in the Appendix and one can follow any one of these table more effectively. The only difference between the tables are due to the property that

$$\begin{aligned}
p(-\infty < Z < K) &= p(-\infty < Z < 0) + p(0 < Z < K) \\
&= 0.5 + p(0 < Z < K)
\end{aligned}$$

Values of left-hand side probabilities for various values of K are listed in Table 1 and that of right-hand side are listed in Table 2 of Appendix. The following set of examples will illustrate this idea of calculting probability of various events concerning a random variable X which has $n\left(\mu, \sigma^2\right)$.

■ Example 4.25

Let X be a normal random variable with $\mu = 5$ and $\sigma = 10$. Find (a) $p(X > 10)$, (b) $p(6 < X < 9)$, (c) $p(X < 8)$, (d) $p(3 < X < 9)$, (e) $p(-5 < X < 15)$, (f) $p(X < 4)$.

Given $X \sim n(5, 100)$ so that the standard normal variable

$$Z = \frac{X-\mu}{\sigma} \sim n(0,1)$$

i.e.,

$$Z = \frac{X-5}{10} \sim n(0,1)$$

We can calculate the probabilities using the standard normal distribution as follows.

(a)
$$\begin{aligned} p(X > 10) &= p\left(\frac{X-5}{10} > \frac{10-5}{10}\right) = p(Z > 0.2) \\ &= 0.5 - 0.0793 \quad \text{(using Table 2)} \\ &= 0.4207. \end{aligned}$$

(b)
$$\begin{aligned} p(6 < X < 9) &= p\left(\frac{6-5}{10} < \frac{X-5}{\sigma} < \frac{9-5}{10}\right) = p(0.1 < Z < 0.4) \\ &= p(0 < Z < 0.4) - p(0 < Z < 0.1) \\ &= 0.1554 - 0.0398 = 0.1156. \end{aligned}$$

(c)
$$\begin{aligned} p(X < 8) &= p\left(\frac{X-5}{10} < \frac{8-5}{10}\right) = p(Z < 0.3) \\ &= 0.5 + p(0 < Z < 0.3) = 0.6179. \end{aligned}$$

(d)
$$\begin{aligned} p(3 < X < 9) &= p\left(\frac{3-5}{10} < \frac{X-5}{10} < \frac{9-5}{10}\right) = p(-0.2 < Z < 0.4) \\ &= p(0 < Z < 0.2) + p(0 < Z < 0.4) \\ &= 0.0793 + 0.1554 = 0.2347. \end{aligned}$$

(e)
$$\begin{aligned} p(-5 < X < 15) &= p\left(\frac{-5-5}{10} < \frac{X-5}{10} < \frac{15-5}{10}\right) = p(-1 < Z < 1) \\ &= 2p(0 < Z < 1) \quad \text{(using symmetry)} \\ &= 2(0.3413) = 0.6826. \end{aligned}$$

(f)
$$\begin{aligned} p(X < 4) &= p\left(\frac{X-5}{10} < \frac{4-5}{10}\right) = p(Z < -0.1) \\ &= 0.5 - p(0 < Z < 0.1) \\ &= 0.5 - 0.0398 = 0.4602. \end{aligned}$$

■

■ Example 4.26

Scores on a certain test, IQ scores are approximately normally distributed with mean $\mu = 100$ and $\sigma = 15$. An individual is selected at random. What is the probability that his score X satisfies $120 < X < 130$?

Given $X \sim n(100, 225)$ so that the standard normal variable $Z = \frac{X-\mu}{\sigma} \sim n(0, 1)$ that is

$$\begin{aligned} Z &= \frac{X-100}{15} \sim n(0,1) \\ p(120 < X < 130) &= p\left(\frac{120-100}{15} < \frac{X-100}{15} < \frac{130-100}{15}\right) \\ &= p(1.33 < Z < 2) \\ &= p(0 < Z < 2) - p(0 < Z < 1.33) \\ &= 0.4772 - 0.4082 = 0.069. \end{aligned}$$

■

■ Example 4.27

If $X \sim n(30, 25)$ find the probabilities that (a) $26 \leq X \leq 40$, (b) $X \geq 45$, (c) $|X - 30| > 6$, since $X \sim n(30, 25)$, $Z = \frac{X-30}{5} \sim n(0, 1)$.

We can calculate the probabilities using the standard normal distribution as follows.

(a) $$\begin{aligned} p(26 \leq X \leq 40) &= p\left(\frac{26-30}{5} \leq \frac{X-30}{5} \leq \frac{40-30}{5}\right) = p\left(-0.8 \leq Z \leq 2\right) \\ &= p(0 \leq Z \leq 0.8) + p(0 \leq Z \leq 2) \\ &= 0.2881 + 0.4772 = 0.7653. \end{aligned}$$

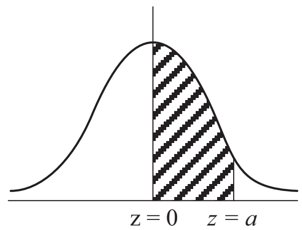
(b) $$\begin{aligned} p(X \geq 45) &= p\left(\frac{X-30}{5} \geq \frac{45-30}{5}\right) = p(Z \geq 3) \\ &= 0.5 - p(0 \leq Z \leq 3) \\ &= 0.5 - 0.4986 = 0.0014. \end{aligned}$$

(c) $$\begin{aligned} p(|X-30| > 6) &= 1 - p(|X-30| \leq 6) = 1 - p(-6 \leq X - 30 \leq 6) \\ &= 1 - p\left(\frac{-6}{5} \leq \frac{X-30}{5} \leq \frac{6}{5}\right) \\ &= 1 - p(-1.2 \leq Z \leq 1.2) \\ &= 1 - 2(0 \leq Z \leq 1.2) \\ &= 1 - 2(0.3849) = 0.2302. \end{aligned}$$

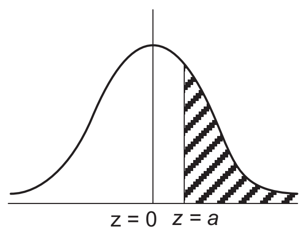
Remark

Let us summarize the various possibilities of computations, we refer Table 2 of Appendix in this attempt. However, calculations can be done referring Table 1 of Appendix also. Assume a standard normal variable 'Z' and two real numbers 'a' and 'b' such that $a > 0$ and $b > 0$.

Case (i): $p(0 < Z < a)$ for various values of a (where $a > 0$) are listed in Table 1.

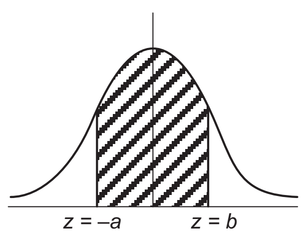


Case (ii): $p(Z > a) = 0.5 - p(0 < Z < a)$.

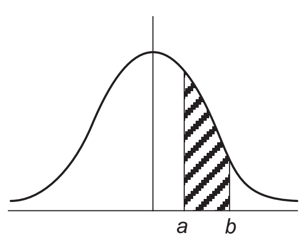


Case (iii): $p(-a < Z < b)$ where $b > 0$

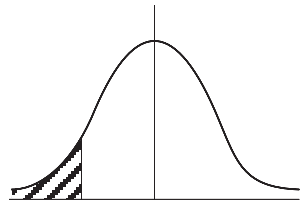
$$
\begin{aligned}
&= p(-a < Z < 0) + p(0 < Z < b) \\
&= p(0 < Z < a) + p(0 < Z < b) \quad \text{(By symmetry)}
\end{aligned}
$$



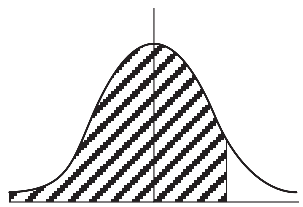
Case (iv): $p(a < Z < b) = p(0 < Z < b) - p(0 < Z < a)$



Case (v): $p(Z < -a) = p(Z > a)$ (By symmetry)
$= 0.5 - p(0 < Z < a)$



Case (vi): $p(Z < a) = 0.5 + p(0 < Z < a)$



Note that Table 2 of Appendix provides the probabilities in the last step of Case (ii) to Case (vi).

■ Example 4.28

Let $X \sim n\left(\mu, \sigma^2\right)$ so that $p(X < 89) = 0.9$ and $p(X < 94) = 0.95$. Find μ and σ^2.

Given $X \sim n\left(\mu, \sigma^2\right)$ and that

$$p(X < 89) = 0.9 \quad \text{and} \quad p(X < 94) = 0.95$$

$$\therefore \quad p\left(\frac{X-\mu}{\sigma} < \frac{89-\mu}{\sigma}\right) = 0.9 \quad \text{and} \quad p\left(\frac{X-\mu}{\sigma} < \frac{94-\mu}{\sigma}\right) = 0.95$$

$$\therefore \quad p(Z < a) = 0.9 \quad \text{and} \quad p(Z < b) = 0.95$$

$$\text{(or)} \quad p(Z < a) = 0.5 + p(0 < Z < a) = 0.9$$

Hence $p(0 < Z < a) = 0.4$ and similarly $p(0 < Z < b) = 0.45$. From Table 2, we have

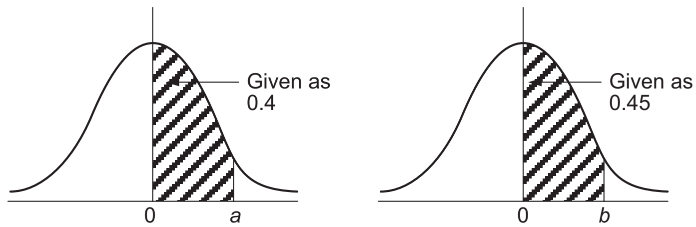
$$a = 1.29 \quad \text{and} \quad b = 1.65$$

$$\therefore \quad \frac{89-\mu}{\sigma} = 1.29 \quad \text{and} \quad \frac{94-\mu}{\sigma} = 1.65$$

$$\text{(or)} \quad 89 - \mu = 1.29\sigma \quad \text{and} \quad 94 - \mu = 1.65\sigma$$

Solving these two equations we get $\sigma = 13.9$ so that $\sigma^2 = 193.2$ and $\mu = 71.1$.

***Note*:** In the above example we find the values of Z for a special proabilities (0.4 and 0.45).



In Table 1, we refer the probabilities closer to 0.4 and 0.45 and infact we assumed $a = 1.29$ and $b = 1.65$. This "closeness" may be varied accordingly the values of a and b also change which influence the values of μ and σ^2 also. Hence, we may get different μ and σ^2. But these "differences" are negligible.

■ Example 4.29

If 8% of the probability for a certain distribution that is $n\left(\mu, \sigma^2\right)$, is below 50 and that 5% is above 90, what are the values of μ and σ^2?

Given that $X \sim n\left(\mu, \sigma^2\right)$ and that

$$p(X \leq 50) = 0.08 \quad \text{and} \quad p(X \geq 90) = 0.05$$

$$p\left(\frac{X-\mu}{\sigma} \leq \frac{50-\mu}{\sigma}\right) = 0.08 \quad \text{and} \quad p\left(\frac{X-\mu}{\sigma} \geq \frac{90-\mu}{\sigma}\right) = 0.05$$

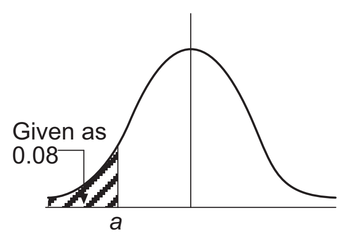
$$\Rightarrow \quad p(Z \leq a) = 0.08 \quad \text{and} \quad p(Z \geq b) = 0.05$$

$$\text{where} \quad a = \frac{50-\mu}{\sigma} \quad \text{and} \quad b = \frac{90-\mu}{\sigma}$$

Also, since $p(Z \leq a) = 0.08 < 0.5$, 'a' should be negative. (Refer Case (v) of summary). From Table 2, $a = -1.41$

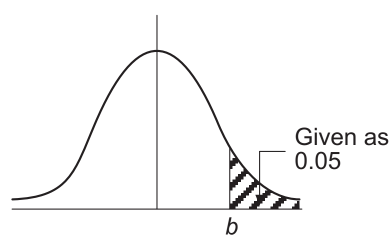
Similarly, $p(Z \geq b) = 0.05$ so that $p(0 < Z < b) = 0.45$ and from Table 1,

$$b = 1.65$$



$$\text{Hence} \quad \frac{50-\mu}{\sigma} = -1.41 \quad \text{and} \quad \frac{90-\mu}{\sigma} = 1.65$$

$$\text{(or)} \quad 50-\mu = -1.41\sigma \quad \text{and} \quad 90-\mu = 1.65\sigma$$



Solving these equations we get $\mu = 68.4$ and $\sigma = 13.1$.

■ Example 4.30

Let $X \sim n(75, 25)$. Find the conditional probability that X is greater than 80 relative to the hypothesis that X is greater than 77.

Given $X \sim n(75, 25)$ so that $\mu = 75$ and $\sigma = 5$.

$$\begin{aligned}
\text{To find} \quad p(X > 80/X > 77) &= \frac{p(X > 80 \cap X > 77)}{p(X > 77)} \\
&= \frac{p(X > 80)}{p(X > 77)} = \frac{p\left(\frac{X-75}{5} > \frac{80-75}{5}\right)}{p\left(\frac{X-75}{5} > \frac{77-75}{5}\right)} \\
&= \frac{p(Z > 1)}{p(Z > 0.4)} = \frac{0.5 - 0.3413}{0.5 - 0.1554} \\
&= 0.461.
\end{aligned}$$

Once, we have introduced some important univariate special distributions in the previous section, now we shall find their characteristics namely mean and variance. Such constants play an important role in statistics. So, we try to derive them either by evaluating the necessary sums or integrals directly or we can work with MGF (if exists). Even though we may use these constants in our subsequent chapters, derivations will merely serve to provide us with experience in the application of the respective mathematical techniques. We try to adopt this approach to derive the constants of discrete and continuous distributions. Further, the most common parameters of some distributions we have discussed are the lower moments, in particular mean (μ) and variance (σ^2).

4.7 MOMENTS OF THE DISTRIBUTIONS

MGF of binomial distribution

Let $X \sim b(n, p)$. Hence, its pmf is

$$p(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x}; & x = 0, 1, 2, \cdots, n \\ 0 & \text{elsewhere} \end{cases}$$

Now, MGF of X is

$$\begin{aligned}
M_X(t) &= E\left(e^{tx}\right) \\
&= \sum_x e^{tx} p(X = x) \\
&= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x} \\
&= \sum_{x=0}^{n} \binom{n}{x} \left(pe^t\right)^x q^{n-x} \\
&= \left(q + pe^t\right)^n
\end{aligned}$$

Mean and variance

Since $M_X(t) = \left(q + pe^t\right)^n$, we have

$$\begin{aligned} M_x'(t) &= n\left(q + pe^t\right)^{n-1} pe^t = np\, e^t \left(q + pe^t\right)^{n-1} \\ \text{and} \quad M_X''(t) &= npe^t \left(q + pe^t\right)^{n-2} \left[q + pe^t + pe^t(n-1)\right] \end{aligned}$$

Putting $t = 0$ in the above equations, we have

$$\begin{aligned} E(X) &= np \quad \text{and} \quad E\left(X^2\right) = np\,[1 + p(n-1)] \\ \text{Hence,} \quad V(X) &= E\left(X^2\right) - E(X)^2 \\ &= np + n^2p^2 - np^2 - n^2p^2 \\ &= np(1-p) = npq \end{aligned}$$

So that mean and variance of a Binomial variate is np and npq respectively.

MGF of the poission distribution

Let $X \sim P(\lambda)$. Hence, its probability distribution is

$$f(x) = \begin{cases} e^{-\lambda}\frac{\lambda^x}{x!} & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Now, MGF of X is

$$\begin{aligned} M_X(t) &= E\left(e^{tx}\right) \\ &= \sum_x e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} e^{-\lambda} \left(\frac{\lambda^x}{x!}\right) \\ &= e^{-\lambda} \sum_{x=0}^{\infty} \frac{\left(\lambda e^t\right)^x}{x!} = e^{-\lambda}\, e^{\lambda e^t} \\ M_X(t) &= e^{\lambda\left(e^t - 1\right)} = e^{-\lambda(1-e^t)} \end{aligned}$$

Mean and variance

If we differentiate $M_X(t)$ twice with respect to 't', we have

$$\begin{aligned} M_X'(t) &= \lambda e^t\, e^{\lambda\left(e^t - 1\right)} \\ \text{and} \quad M_X''(t) &= \lambda e^t\, e^{\lambda\left(e^t - 1\right)} \left(1 + \lambda e^t\right) \end{aligned}$$

So that, putting $t = 0$ in the above equations, we have $E(X) = \lambda$ and $E(X^2) = \lambda(1 + \lambda)$. Thus, $V(X) = E(X^2) - E(X)^2 = \lambda(1 + \lambda) - \lambda^2 = \lambda$. Hence, the mean and the variance of the Poisson distributions are $\mu = \lambda$ and $\sigma^2 = \lambda$ respectively.

MGF of the negative binomial distribution

Let $X \sim nb(x : r, p)$. Hence, its probability distribution is

$$f(x) = \begin{cases} \binom{x+r-1}{r-1} p^r q^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

MGF of X

$$\begin{aligned} M_X(t) &= E\left(e^{tx}\right) \\ &= \sum_x e^{tx} f(x) = \sum_{x=0}^{\infty} e^{tx} \binom{x+r-1}{r-1} p^r q^x \\ &= p^r \sum_{x=0}^{\infty} \binom{x+r-1}{r-1} \left(qe^t\right)^x \\ &= p^r \left[1 + \binom{r}{r-1} qe^t + \binom{r+1}{r-1} \left(qe^t\right)^2 + \binom{r+2}{r-1} \left(qe^t\right)^3 + \cdots\right] \\ &= p^r \left[1 + \binom{r}{1} \left(qe^t\right) + \binom{r+1}{2} \left(qe^t\right)^2 + \binom{r+2}{3} \left(qe^t\right)^3 + \cdots\right] \end{aligned}$$

Since $\binom{n}{r} = \binom{n}{n-r}$

$$\begin{aligned} &= p^r \left[1 + r\left(qe^t\right) + \frac{r(r+1)}{2!} \left(qe^t\right)^2 + \frac{r(r+1)(r+2)}{3!} \left(qe^t\right)^3 + \cdots\right] \\ &= p^r \left(2 - qe^t\right)^{-r} \end{aligned}$$

Hence, $$M_X(t) = \left(\frac{p}{1-qe^t}\right)^r$$

If we differentiate $M_X(t)$ twice with respect to t, we get

$$\begin{aligned} M'_X(t) &= p^r \left(\frac{d}{dt}\right) \left[\frac{1}{1-qe^t}\right]^r \\ &= p^r (-r) \left(1-qe^t\right)^{-r-1} \left(-qe^t\right) \\ &= rp^r q \left[\frac{e^t}{(1-qe^t)^{r+1}}\right] \quad \text{or} \quad rp^r q\, e^t \left(1-qe^t\right)^{-r-1} \\ M''_X(t) &= rp^r q \left[e^t \left(1-qe^t\right)^{-r-1} + (-r-1)\left(1-qe^t\right)^{-r-2} \left(-qe^t\right) \cdot e^t\right] \\ &= rp^r q \left(1-qe^t\right)^{-r-2} e^t \left[1 - qe^t + (r+1)qe^t\right] \\ &= rp^r q\, e^t \frac{\left(1+rqe^t\right)}{(1-qe^t)^{r+2}} \end{aligned}$$

and upon substituting $t = 0$, we get

$$\begin{aligned}
\mu = E(X) &= \left(M'_X(t)\right)_{t=0} \\
&= rp^r q\left[\frac{1}{(1-q)^{r+1}}\right] = rp^r q \cdot \frac{1}{p^{r+1}} \\
&= \frac{rq}{p} \quad \text{and} \\
M'_2 = E\left(X^2\right) &= \left[M'_X(t)\right]_{t=0} \\
&= rp^r q\left[\frac{1+rq}{(1-q)^{r+2}}\right] = rp^r q\left[\frac{1+rq}{p^{r+2}}\right] = \frac{rq(1+rq)}{p^2}
\end{aligned}$$

Hence, variance of X is

$$\begin{aligned}
\sigma^2 &= E\left(X^2\right) - E(X)^2 \\
&= \frac{rq(1+rq)}{p^2} - \frac{r^2q^2}{p^2} = \frac{rq}{p^2}
\end{aligned}$$

MGF of the geometric distribution

Let $X \sim Ge(p)$. Hence, the probability distribution of X is

$$f(x) = \begin{cases} p(1-p)^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Now, MGF of X is

$$\begin{aligned}
M_X(t) &= E\left(e^{tx}\right) \\
&= \sum_x e^{tx} f(x) \\
&= \sum_{x=0}^{\infty} e^{tx} p(1-p)^x = p\sum_{x=0}^{\infty} e^{tx} q^x \\
&= p\left[1 + qe^t + \left(qe^t\right)^2 + \left(qe^t\right)^3 + \cdots\right] \\
M_X(t) &= \frac{p}{1-q^{e^t}}
\end{aligned}$$

If we differentiate $M_X(t)$ twice with respect to 't', we have

$$M'_X(t) = \frac{pqe^t}{(1-qe^t)^2} \quad \text{and} \quad M''_X(t) = pqe^t\left(\frac{1+qe^t}{(1-qe^t)^3}\right)$$

So that putting $t = 0$ in the above equations we get mean of the Geometric distribution as

$$\mu = E(X) = \left[M'_X(t)\right]_{t=0} = \frac{pq}{(1-q)^2} = \frac{pq}{p^2} = \frac{q}{p}$$

$$\text{and variance} \quad \sigma^2 = V(X) = E(X^2) - E(X)^2 = \left[M_X''(t)\right]_{t=0} - E(X)^2$$
$$= \frac{pq(1+q)}{(1-q)^3} - \frac{q^2}{p^2} = \frac{q+q^2-q^2}{p^2} = \frac{q}{p^2}$$

***Note*:** Some authors define the geometric distribution as $f(x) = pq^{x-1}, x = 1, 2, \cdots$. In such a case, its MGF is $\frac{pe^t}{1-qe^t}$ and hence its mean is $\frac{1}{p}$ and variance is $\frac{q}{p^2}$.

MGF of the discrete uniform distribution

Let X have a uniform distribution. Then, its probability distribution is

$$f(x) = \begin{cases} \frac{1}{K} & x = 1, 2, \cdots K \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned}
\text{Now,} \quad M_X(t) &= E\left(e^{tx}\right) \\
&= \sum_x e^{tx} f(x) \\
&= \sum_{x=1}^{K} e^{tx} \cdot \frac{1}{K} \\
&= \frac{1}{K}\left[e^t + e^{2t} + e^{3t} + \cdots + e^{Kt}\right] \qquad (1) \\
&= \frac{1}{K}e^t\left[1 + \left(e^t\right)^2 + (e^t)^3 + \cdots + (e^t)^{K-1}\right] \\
&= \frac{1}{K}e^t\left[\frac{1-\left(e^t\right)^K}{1-e^t}\right] \\
M_X(t) &= \frac{1}{K}e^t\left[\frac{1-e^{Kt}}{1-e^t}\right] \qquad (2)
\end{aligned}$$

Instead of differentiating $M_X(t)$ in equation (2) with respect to t, let us differentiate $M_X(t)$ in equation (1) twice with respect to t, we have

$$\begin{aligned}
M_X'(t) &= \frac{1}{K}\left[e^t + 2e^{2t} + 3e^{3t} + \cdots + Ke^{Kt}\right] \\
M_X''(t) &= \frac{1}{K}\left[e^t + 2^2e^{2t} + 3^2e^{3t} + \cdots + K^2e^{Kt}\right]
\end{aligned}$$

and upon substituting $t = 0$, we get

$$\begin{aligned}
\mu_1' = E(X) &= \left[M_X'(t)\right]_{t=0} \\
&= \frac{1}{K}[1 + 2 + 3 + \cdots + K]
\end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{K}\left[\frac{K(K+1)}{2}\right] \\
&= \frac{K+1}{2} \\
\mu_2' = E(X^2) &= \left[M_X''(t)\right]_{t=0} \\
&= \frac{1}{K}\left[1+2^2+3^2+\cdots+K^2\right] \\
&= \frac{1}{K}\left[\frac{K(K+1)(2K+1)}{6}\right] \\
&= \frac{(K+1)(2K+1)}{6}
\end{aligned}
$$

Hence, variance of X is

$$
\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \frac{(K+1)(2K+1)}{6} - \left(\frac{K+1}{2}\right)^2 \\
&= \frac{\left(2K^2+3K+1\right)}{6} - \frac{\left(K^2+2K+1\right)}{4} \\
&= \frac{K^2-1}{12}
\end{aligned}
$$

Mean and variance of the hyper geometric distribution

Let $X \sim hy(x; N, M, n)$. Then, the probability distribution of X is

$$
f(x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & x = 0, 1, \cdots, n \\ 0 & \text{elsewhere} \end{cases}
$$

$$
\begin{aligned}
E(x) &= \sum_x x f(x) = \sum_{x=0}^{n} x \binom{M}{x}\binom{N-M}{n-x} \Big/ \binom{N}{n} \\
&= 0 + \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} x \cdot \left(\frac{M}{x}\right)\binom{M-1}{x-1}\binom{N-M}{n-x} \quad \text{(when } x = 0\text{, first item vanishes)} \\
&= M \cdot \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} \binom{M-1}{x-1}\binom{N-1-M+1}{n-1-x+1} \\
&= M \cdot \frac{1}{\binom{N}{n}} \sum_{x=1}^{n} \binom{M-1}{x-1}\binom{N-1-M+1}{n-1-(x-1)}
\end{aligned}
$$

Let $x-1=y$ so that $y=0$ to $n-1$

$$= M\frac{1}{\binom{N}{n}}\sum_{y=0}^{n-1}\binom{M-1}{y}\binom{N-1-M+1}{n-1-y}$$

$$= M\cdot\frac{1}{\binom{N}{n}}\binom{M-1+N-1-M+1}{n-1}$$

$$= M\frac{1}{\binom{N}{n}}\binom{N-1}{n-1} = M\frac{1}{\frac{N}{n}\binom{N-1}{n-1}}\binom{N-1}{n-1}$$

$$\therefore \quad E(X) = \frac{nM}{N}$$

$$E(X^2) = \sum_x x^2 f(x) = \sum_{x=0}^{n} x^2\binom{M}{x}\binom{N-M}{n-x}\Big/\binom{N}{n}$$

$$= \frac{1}{\binom{N}{n}}\sum_{x=1}^{n} x^2\left(\frac{M}{x}\right)\binom{M-1}{x-1}\binom{N-M}{n-x}$$

$$= M\frac{1}{\binom{N}{n}}\sum_{x}^{n}(x-1+1)\binom{M-1}{x-1}\binom{N-M}{n-x}$$

$$= M\cdot\frac{1}{\binom{N}{n}}\left[\sum_{x=1}^{n}(x-1)\binom{M-1}{x-1}\binom{N-M}{n-x}+\sum_{x=1}^{n}\binom{M-1}{x-1}\binom{N-M}{n-x}\right]$$

$$= M\cdot\frac{1}{\binom{N}{n}}\left[\sum_{x=2}^{n}\left(\frac{M-1}{x-1}\right)\cdot(x-1)\binom{M-2}{x-2}\binom{N-M}{n-x}+\sum_{x=1}^{n}\binom{M-1}{x-1}\binom{N-M}{n-x}\right]$$

$$= M\cdot\frac{1}{\binom{N}{n}}\left[(M-1)\sum_{x=1}^{n}\binom{M-2}{x-2}\binom{N-2-M+2}{n-2-x+2}+\sum_{x=1}^{n}\binom{M-1}{x-1}\binom{N-1-M+1}{n-1-(x-1)}\right]$$

Let $y = x - 2$ so that $y = 0$ to $n - 2$ and $z = x - 1$, so that $z = 0$ to $n - 1$.

$$
\begin{aligned}
&= M\frac{1}{\binom{N}{n}}\left[(M-1)\sum_{y=0}^{n-2}\binom{M-2}{y}\binom{N-2-M+2}{n-2-y}+\right.\\
&\qquad\left.\sum_{z=0}^{n-1}\binom{M-1}{Z}\binom{N-1-M+1}{n-1-Z}\right]\\
&= M\frac{1}{\binom{N}{n}}\left[(M-1)\binom{N-2}{n-2}+\binom{N-1}{n-1}\right]\\
&= M(M-1)\frac{1}{\binom{N}{n}}\binom{N-2}{n-2}+M\frac{1}{\binom{N}{n}}\binom{N-1}{n-1}\\
&= M\cdot(M-1)\frac{1}{\frac{N}{n}\frac{N-1}{n-1}}+M\cdot\frac{1}{\frac{N}{n}}\\
&= \frac{M(M-1)\,n(n-1)}{N(N-1)}+\frac{nM}{N}
\end{aligned}
$$

Hence,

$$
\begin{aligned}
V(X) &= E(X^2)-E(X)^2\\
&= \frac{M(M-1)n(n-1)}{N(N-1)}+\frac{nM}{N}-\frac{n^2M^2}{N^2}\\
&= \frac{nM}{N}\left[\frac{(M-1)(n-1)}{N-1}+1-\frac{nM}{N}\right]\\
&= \frac{nM}{N}\left[\frac{(M-1)(n-1)N+N(N-1)-nM(N-1)}{N(N-1)}\right]\\
&= \frac{nM}{N}\left[\frac{MNn-MN-nN+N+N^2-N-nMN+nM}{N(N-1)}\right]\\
&= \frac{nM}{N}\left[\frac{N(N-M)-n(N-M)}{N(N-1)}\right]\\
&= \frac{nM}{N}\left[\frac{(N-m)(N-n)}{N(N-1)}\right]
\end{aligned}
$$

MGF of uniform distribution

Let $X \sim u(a, b)$. Hence, its probability distribution is

$$
f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{elsewhere} \end{cases}
$$

Hence, MGF of X is

$$\begin{aligned}
M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx = \int_a^b e^{tx} \frac{1}{b-a} dx \\
&= \frac{1}{b-a}\left(\frac{e^{tx}}{t}\right)_a^b = \frac{1}{b-a}\left(\frac{e^{bt}-e^{at}}{t}\right) \\
&= \frac{1}{b-a}\left(\frac{1}{t}\right)\left[\left(1+\frac{bt}{1!}+\frac{b^2t^2}{2!}+\frac{b^3t^3}{3!}+\cdots\right)-\right. \\
&\qquad \left.\left(1+\frac{at}{1!}+\frac{a^2t^2}{2!}+\frac{a^3t^3}{3!}+\cdots\right)\right] \\
&= \frac{1}{b-a}\left[(b-a)+\left(\frac{b^2-a^2}{2}\right)t+\left(\frac{b^3-a^3}{3}\right)\frac{t^2}{2!}+\cdots\right]
\end{aligned}$$

$$\begin{aligned}
\mu_1' &= E(X) = \text{Coefficient of } \left(\frac{t}{1!}\right) \text{ in } M_X(t) = \frac{b^2-a^2}{2(b-a)} \\
\mu_2' &= E(X^2) = \text{Coefficient of } \left(\frac{t^2}{2!}\right) \text{ in } M_X(t) = \frac{b^3-a^3}{3(b-a)}
\end{aligned}$$

Hence, mean of X is

$$\mu = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

and

$$E(X^2) = \frac{a^2+ab+b^2}{3}$$

so that variance of X is

$$\begin{aligned}
V(X) &= E(X^2) - E(X)^2 \\
&= \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} \\
&= \frac{(b-a)^2}{12}
\end{aligned}$$

MGF of normal distribution

Let $X \sim n(\mu, \sigma^2)$. Hence, its probability distribution is

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty < x < \infty$$

Hence MGF of X is

$$\begin{aligned}
M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)\, dx \\
&= \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}\, dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}+tx}\, dx \qquad (1)
\end{aligned}$$

Consider

$$\begin{aligned}
\frac{(x-\mu)^2}{2\sigma^2} - tx &= \frac{1}{2\sigma^2}\left[(x-\mu)^2 - 2\sigma^2 tx\right] \\
&= \frac{1}{2\sigma^2}\left[x^2 + \mu^2 - 2\mu x - 2\sigma^2 tx\right] \\
&= \frac{1}{2\sigma^2}\left[x^2 - 2x\left(\sigma^2 t + \mu\right) + \mu^2\right] \\
&= \frac{1}{2\sigma^2}\left[\left(x - \left(\sigma^2 t + \mu\right)\right)^2 - \left(\sigma^2 t + \mu\right)^2 + \mu^2\right] \\
&= \frac{1}{2\sigma^2}\left[\left(x - \left(\sigma^2 t + \mu\right)\right)^2 - \sigma^2\left(\sigma^2 t^2 + 2\mu t\right)\right] \quad \text{(power)} \\
&= \frac{\left(x - \left(\sigma^2 t + \mu\right)\right)^2}{2\sigma^2} - \frac{1}{2}\left(\sigma^2 t^2 + 2\mu t\right)
\end{aligned}$$

Hence equation (1) becomes

$$\begin{aligned}
M_X(t) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\left[x-\left(\sigma^2 t+\mu\right)\right]^2}{2\sigma^2}} e^{\frac{1}{2}\left(\sigma^2 t^2 + 2\mu t\right)} dx \\
&= \frac{1}{\sigma\sqrt{2\pi}} e^{\mu t + \frac{\sigma^2 t^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{\left[x-\left(\sigma^2 t+\mu\right)\right]^2}{\sqrt{2}\sigma}} dx
\end{aligned}$$

Now, if we substitute $Z = \left(\frac{x-(\sigma^2 t+\mu)}{\sqrt{2}\sigma}\right)$, then $dz = \left(\frac{1}{\sqrt{2}\sigma}\right) dx$ so that

$$\begin{aligned}
M_X(t) &= \frac{1}{\sqrt{\pi}} e^{\mu t + \frac{\sigma^2 t^2}{2}} \int_{-\infty}^{\infty} e^{-Z^2} dz \quad \text{(using Gamma Integral)} \\
&= \frac{1}{\sqrt{\pi}} e^{\mu t + \frac{\sigma^2 t^2}{2}} \sqrt{\pi} \\
M_X(t) &= e^{\mu t + \frac{\sigma^2 t^2}{2}}
\end{aligned}$$

Differentiate $M_X(t)$ twice with respect to t, we have

$$\begin{aligned} M_X'(t) &= e^{\mu t+\frac{\sigma^2 t^2}{2}}\left(\mu+\sigma^2 t\right) \\ M_X''(t) &= e^{\mu t+\frac{\sigma^2 t^2}{2}}\left[\left(\mu+\sigma^2 t\right)^2+\sigma^2\right] \end{aligned}$$

and upon substituting $t = 0$, we have

$$\begin{aligned} E(X) &= \left[M_X'(t)\right]_{t=0} = \mu \\ E\left(X^2\right) &= \mu^2+\sigma^2 \end{aligned}$$

so that variance of X is

$$\begin{aligned} V(X) = E(X^2) - [E(X)]^2) &= \mu^2+\sigma^2-\mu^2 \\ &= \sigma^2 \end{aligned}$$

***Note*:** From the MGF of normal distribution, we can easily find the MGF of standard normal distribution as $M_X(t) = e^{t^2/2}$ and hence mean 0 and variance 1.

MGF of gamma distribution

Let $X \sim G(x; \alpha, \beta)$. Its probability distribution is

$$f(x) = \begin{cases} \frac{1}{\lceil\alpha\ \beta^\alpha} x^{\alpha-1}\ e^{-x/\beta} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Its MGF is

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)dx \\ &= \int_0^{\infty} e^{tx}\frac{1}{\lceil\alpha\ \beta^\alpha} x^{\alpha-1} e^{x/\beta}dx \\ &= \frac{1}{\lceil\alpha\ \beta^\alpha}\int_0^{\infty} x^{\alpha-1} e^{-x\left(\frac{1}{\beta}-t\right)}dx \end{aligned}$$

Substituting $x\left(\frac{1}{\beta}-t\right) = z$, we have $x, = dx\left(\frac{1}{\beta}-t\right)$. Hence, this integral becomes,

$$\begin{aligned} M_X(t) &= \frac{1}{\lceil\alpha\ \beta^\alpha}\int_0^{\infty} Z^{\alpha-1}e^{-Z}\left[\frac{1}{\left(\frac{1}{\beta}-t\right)^\alpha}\right]dz \\ &= \frac{1}{\beta^\alpha}\cdot\frac{1}{\left(\frac{1}{\beta}-t\right)^\alpha}\cdot\frac{1}{\lceil\alpha}\int_0^{\infty} Z^{\alpha-1}e^{-z}\,dz \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{\beta^\alpha}\frac{\beta^\alpha}{(1-\beta t)^\alpha}\cdot\frac{1}{\lceil\overline{\alpha}}\lceil\overline{\alpha} \\
\therefore \qquad M_X(t) &= \frac{1}{(1-\beta t)^\alpha}
\end{aligned}
$$

Differentiating $M_X(t)$,twice with respect to t,

$$M'_X(t) = \frac{\alpha\beta}{(1-\beta t)^{\alpha+1}} \quad \text{and} \quad M''_X(t) = \alpha\beta^2\frac{\alpha+1}{(1-\beta t)^{\alpha+2}}$$

and upon substituting $t = 0$, we have

$$
\begin{aligned}
E(X) &= \left[M'_X(t)\right]_{t=0} = \alpha\beta \\
E\left(X^2\right) &= \left[M''_X(t)\right]_{t=0} = \alpha\beta^2(\alpha+1)
\end{aligned}
$$

Hence, variance of X is

$$
\begin{aligned}
V(X) &= E\left(X^2\right) - E(X)^2 \\
&= \alpha\beta^2(\alpha+1) - \alpha^2\beta^2 \\
&= \alpha\beta^2
\end{aligned}
$$

***Note*:** From the MGF of gamma distribution we can find the MGF of the special case of gamma distribution in which $\alpha = \frac{r}{2}$ and $\beta = 2$ (we defined it as chi-square distribution). Let $X \sim \chi^2(r)$.

Hence $\qquad M_X(t) = \dfrac{1}{(1-2t)^{r/2}}$ and

Mean $\qquad = \dfrac{r}{2}\cdot 2 = r$ and variance $= \frac{r}{2}\cdot 2^2 = 2r$.

The other special case of gamma distribution in which $\alpha = 1$ and $\beta = \frac{1}{\lambda}$ (we defined it as Exponential distribution). If $X \sim Expo(\lambda)$, then its $M_X(t) = \frac{\lambda}{\lambda - t}$ mean $= \frac{1}{\lambda}$ and variance $= \frac{1}{\lambda^2}$. Our next derivation is to obtain these values of exponential distribution using a direct integration.

MGF of exponential distribution

Let $X \sim Expo(\lambda)$. Its probability distribution is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \lambda > 0; x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Its MGF is

$$
\begin{aligned}
M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx \\
&= \int_{0}^{\infty} e^{tx}(\lambda)e^{-\lambda x}\,dx
\end{aligned}
$$

$$= \lambda \int_0^\infty e^{-x(\lambda-t)}\, dx$$

$$= \lambda \left(\frac{e^{-x(\lambda-t)}}{-(\lambda-t)} \right)_0^\infty$$

$$M_X(t) = \frac{\lambda}{\lambda - t}$$

Differentiate $M_X(t)$ twice with respect to t,

$$M'_X(t) = \frac{\lambda}{(\lambda-t)^2} \quad \text{and} \quad M''_X(t) = \frac{2\lambda}{(\lambda-t)^3}$$

and upon substituting $t = 0$, we have

$$E(X) = \left[M'_X(t)\right]_{t=0} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$E\left(X^2\right) = \left[M''_X(t)\right]_{t=0} = \frac{2\lambda}{\lambda^3} = \frac{2}{\lambda^2}$$

Hence, variance of X is $V(X) = E\left(X^2\right) - E(X)^2$

$$= \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Moments of weibull distribution

Let $X \sim We(x; \alpha, \beta)$. Its pdf is

$$f(x) = \begin{cases} (\alpha\beta)\, x^{\beta-1}\, e^{-\alpha x^\beta} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$\mu'_r = E(X^r) = \int_{-\infty}^{\infty} x^r\ f(x)\, dx$$

$$= \int_0^\infty x^r (\alpha\beta)\, x^{\beta-1}\, e^{-\alpha x^\beta}\, dx$$

Substituting $\alpha x^\beta = z$, we have $dz = (\alpha\beta)x^{\beta-1}\, dx$

$$\therefore \quad E\left(X^r\right) = \int_0^\infty \left[\left(\frac{z}{\alpha}\right)^{1/\beta}\right]^r e^{-z} dz$$

$$= \frac{1}{\alpha^{r/\beta}} \int_0^\infty z^{\left(\frac{r}{\beta}+1\right)-1} e^{-z}\, dz$$

$$= \frac{1}{\alpha^{r/\beta}} \Gamma\left(1+\frac{r}{\beta}\right)$$

If we take $r = 1$ and $r = 2$, we have

$$\text{mean} \qquad \mu = E(X) = \frac{1}{\alpha^{1/\beta}} \Gamma\left(1+\frac{1}{\beta}\right) \qquad \text{and} \qquad E\left(X^2\right) = \frac{1}{\alpha^{2/\beta}} \Gamma\left(1+\frac{2}{\beta}\right)$$

so that variance of X is

$$\begin{aligned} V(X) &= E\left(X^2\right) - E(X)^2 \\ &= \frac{1}{\alpha^{2/\beta}} \Gamma\left(1+\frac{2}{\beta}\right) - \frac{1}{\alpha^{2/\beta}} \left(\Gamma\left(1+\frac{2}{\beta}\right)\right)^2 \\ &= \frac{1}{\alpha^{2/\beta}} \left[\Gamma\left(1+\frac{2}{\beta}\right) - \left(\Gamma\left(1+\frac{1}{\beta}\right)\right)^2\right] \end{aligned}$$

Moments of beta distribution

Let $X \sim Be(x; \alpha, \beta)$ than its probability distribution is

$$f(x) = \begin{cases} \frac{1}{\beta(\alpha,\beta)} x^{\alpha-1(1-x)^{\beta-1}} & 0 < x < 1 \\ 0 & \text{elsewhere} \end{cases}$$

Let us find $E\left(X^r\right)$

$$\begin{aligned} E\left(X^r\right) &= \int_{-\infty}^{\infty} x^r\, f(x)\, dx = \int_0^1 x^r \cdot \frac{1}{\beta(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}\, dx \\ &= \frac{1}{\beta(\alpha,\beta)} \int_0^1 x^{r+\alpha-1}(1-x)^{\beta-1}\, dx = \frac{1}{\beta(\alpha,\beta)}\, \beta(r+\alpha, \beta) \\ &= \frac{1}{\frac{\Gamma(\alpha)\,\Gamma(\beta)}{\Gamma(\alpha+\beta)}} \frac{\Gamma(r+a)\,\Gamma(\beta)}{\Gamma(r+\alpha+\beta)} = \frac{\Gamma(r+a)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(r+\alpha+\beta)} \end{aligned}$$

Hence, the mean of X

$$E(X) = \frac{\Gamma(1+\alpha)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,\Gamma(1+\alpha+\beta)} = \frac{\alpha\,\Gamma(\alpha)\,\Gamma(\alpha+\beta)}{\Gamma(\alpha)\,(\alpha+\beta)\,\Gamma(\alpha+\beta)}$$

$$= \frac{\alpha}{\alpha+\beta}$$

$$\text{and} \qquad E(X^2) = \frac{\sqrt{2+\alpha}\sqrt{\alpha+\beta}}{\sqrt{\alpha}\sqrt{2+\alpha+\beta}} = \frac{\alpha(1+\alpha)\sqrt{\alpha}\sqrt{\alpha+\beta}}{\sqrt{\alpha}(\alpha+\beta+1)(\alpha+\beta)\sqrt{\alpha+\beta}}$$

$$= \frac{\alpha(1+\alpha)}{(\alpha+\beta+1)(\alpha+\beta)}$$

Hence, variance of X is

$$\begin{aligned} V(X) &= E\left(X^2\right) - E(X)^2 \\ &= \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \\ &= \frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta+1)(\alpha+\beta)^2} \\ &= \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2} \end{aligned}$$

Next are consider some properties of these special distributions. One important property is additive property or reproductive property that is, sum of independent random variables which follow a special distribution is also a random variable and follows the same distribution. However, not all distributions possess this property and even it does, we may have certain constraints on the values of the parameters of the distribution. In proving this property (for two variables) or some basic examples in this section we utilise the MGF properties, in particular its uniqueness. In fact we will do more on this property in our future chapters. Also, some more properties related to specific distributions will be discussed. We start with the additive property of a binomial distribution.

4.8 CERTAIN PROPERTIES

Property 1: The binomial distribution possesses the additive property if $p_1 = p_2 = p$.

Let $X \sim b(n_1, p)$ and $Y \sim b(n_2, p)$ be two independent random variables. Their MGFs are

$$M_X(t) = \left(q + pe^t\right)^{n_1} \qquad \text{and} \qquad M_Y(t) = \left(q + pe^t\right)^{n_2}$$

$$\begin{aligned} \therefore \qquad M_{X+Y}(t) &= M_X(t) \cdot M_Y(t) \\ &= \left(q + pe^t\right)^{n_1+n_2} \end{aligned}$$

Hence $X + Y \sim b\,(n_1 + n_2, p)$ (i.e.,) $X + Y$ follows binomial distribution with parameters $n_1 + n_2$ and p.

Property 2: Sum of independent Poisson variables is also a Poisson variable.

Let $X \sim P(\lambda_1)$ and $Y \sim P(\lambda_2)$ be two independent Poission variables.

$$\therefore \qquad M_X(t) = e^{\lambda_1\left(e^t-1\right)} \quad \text{and} \quad M_Y(t) = e^{\lambda_2\left(e^t-1\right)}$$

$$\text{Now,} \quad \begin{aligned} M_{X+Y}(t) &= M_X(t) \cdot M_Y(t) \\ &= e^{\lambda_1(e^t-1)} \cdot e^{\lambda_2(e^t-1)} \\ &= e^{(\lambda_1+\lambda_2)(e^t-1)} \\ &= e^{\lambda(e^t-1)} \quad \text{where} \quad \lambda = \lambda_1 + \lambda_2 \end{aligned}$$

Hence $X + Y \sim P\ (\lambda_1 + \lambda_2)$ i.e., $X + Y$ is a Poisson variable with parameter $\lambda = \lambda_1 + \lambda_2$.

Property 3: The sum of two independent normal variables is also a normal variable.

Let $X \sim n(\mu_1, \sigma_1^2)$ and $Y \sim n\left(\mu_2, \sigma_2^2\right)$. Then, we have

$$M_X(t) = e^{\mu_1 t + \sigma_1^2 t^2/2} \quad \text{and} \quad M_Y(t) = e^{\mu_2 t + \sigma_2^2 t^2/2}$$

$$\text{Now,} \quad \begin{aligned} M_{X+Y}(t) &= M_X(t) \cdot M_Y(t) \\ &= e^{\mu_1 t + \sigma_1^2 t^2/2} \cdot e^{\mu_2 t + \sigma_2^2 t^2/2} \\ &= e^{(\mu_1+\mu_2)t + t^2/2\left(\sigma_1^2+\sigma_2^2\right)} \\ &= e^{\mu t + \sigma^2 t^2/2} \end{aligned}$$

where $\mu = \mu_1 + \mu_2$ and $\sigma^2 = \sigma_1^2 + \sigma_2^2$.

Hence $x + Y$ is a normal variable with parameters μ and σ, that is

$$X + Y \sim n\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$$

Property 4: The sum of two independent gamma variables is also a gamma variable.

Let X and Y be two gamma variables with parameters α_1 and β and α_2 and β respectively, that is $X \sim G(x; \alpha_1, \beta)$ and $Y \sim G(y; \alpha_2, \beta)$.

$$\therefore \quad \text{MGF of } X \text{ is } M_X(t) = \frac{1}{(1-\beta t)^{\alpha_1}} \quad \text{and} \quad M_Y(t) = \frac{1}{(1-\beta t)^{\alpha_2}} \text{ so that}$$

$$\begin{aligned} M_{X+Y}(t) &= M_X(t) \cdot M_Y(t) \\ &= \frac{1}{(1-\beta t)^{\alpha_1}} \cdot \frac{1}{(1-\beta t)^{\alpha_2}} \\ &= \frac{1}{(1-\beta t)^{\alpha_1+\alpha_2}} \end{aligned}$$

which is the MGF of a gamma variable with parameters $\alpha_1 + \alpha_2$ and β (or) $Z = X + Y \sim G(Z; \alpha, \beta)$ where $\alpha = \alpha_1 + \alpha_2$.

Property 5: If X and Y are independent Poisson variables the conditional distribution of X given $X + Y$ is binomial.

Let X and Y be independent Poisson variables with parameters λ_1 and λ_2 respectively. Then, $X + Y$ is also a Poisson variable with parameter $\lambda = \lambda_1 + \lambda_2$.

$$
\begin{aligned}
\text{Consider}\quad p(X/X+Y) &= p(X=r/X+Y=n)\\
&= \frac{p(X=r\cap\ X+Y=n)}{p(X+Y=n)}\\
&= \frac{p(X=r\cap\ Y=n-r)}{p(X+Y=n)}\\
&= \frac{p(X=r)\cdot p(Y=n-r)}{p(X+Y=n)} \quad \text{(since } X \text{ and } Y \text{ are independent)}\\
&= \frac{e^{-\lambda_1}\left(\frac{\lambda_1^r}{r!}\right)e^{-\lambda_2}\left(\frac{\lambda_2^{n-r}}{(n-r)!}\right)}{e^{-\lambda}\left(\frac{\lambda^n}{n!}\right)}\\
&= \left(\frac{n!}{r!(n-r)!}\right)\left(\frac{e^{-(\lambda_1+\lambda_2)}}{e^{-\lambda}}\right)\left(\frac{\lambda_1^r\lambda_2^{n-r}}{(\lambda_1+\lambda_2)^n}\right) \quad \text{(since } \lambda=\lambda_1+\lambda_2)\\
&= \binom{n}{r}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^r\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-r}\\
&= \binom{n}{r}p^r(1-p)^{n-r} \qquad \text{where}\quad p=\frac{\lambda_1}{\lambda_1+\lambda_2} \text{ so that}\\
&\qquad\qquad\qquad\qquad q=1-p=\frac{\lambda_2}{\lambda_1+\lambda_2}\\
&= \binom{n}{r}p^rq^{n-r} \qquad r=0,1,2,\cdots,n
\end{aligned}
$$

which is the pdf of a binomial distribution. Hence, the conditional distribution of X given $X+Y=n$ is a binomial distribution with parameters n and $p=\frac{\lambda_1}{\lambda_1+\lambda_2}$.

Property 6: Let the two independent random variables X and Y have the geometric distribution. Then the conditional distribution of X given $X+Y=n$ is uniform.

Since X and Y are independent geometric variables, we have

$$f_X(x)=f_Y(y)=pq^r \quad r=0,1,2,\cdots$$

$$
\begin{aligned}
\text{Now,}\quad p(X=r/X+Y=n) &= \frac{p(X=r\ \cap\ X+Y=n)}{p(X+Y=n)}\\
&= \frac{p(X=r\cap Y=n-r)}{p(X+Y=n)}\\
&= \frac{p(X=r)\cdot p(Y=n-r)}{\sum_{s=0}^{n}[p\,(X_1=s)\cap(X_2=n-s)]}\\
&\qquad \text{(since } X \text{ and } Y \text{ are independent)}\\
&= \frac{p(X=r)p(Y=n-r)}{\sum_{s=0}^{n}[p\,(X_1=s)\,p\,(X_2=n-s)]}
\end{aligned}
$$

$$= \frac{(pq^r)\left(pq^{n-r}\right)}{\sum\limits_{s=0}^{n}[pq^s \cdot pq^{n-s}]} = \frac{p^2q^n}{\sum\limits_{s=0}^{n} p^2q^n}$$

$$= \frac{p^2q^n}{p^2q^n(n+1)} \quad \text{(since, the summation is taken from } s = 0 \text{ to } n \text{ it has } n+1 \text{ terms)}$$

$$= \frac{1}{n+1} \quad r = 0, 1, 2, \cdots, n$$

which is the pmf of a discrete uniform distribution. Hence, the conditional distribution of X given $X + Y = n$ is uniform.

Property 7: If X has a geometric distribution with parameter p, then $p[X \geq K + t/X \geq K] = p[X \geq t]$.

Given that $X \sim Ge(p)$ its pmf is

$$f(x) = \begin{cases} pq^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Now,

$$\begin{aligned} p\,[X \geq K] &= p\,(X = K) + p\,(X = k+1) + \cdots \\ &= pq^K + pq^{K+1} + pq^{K+2} + \cdots \\ &= pq^K\left[1 + q + q^2 + q^3 + \cdots\right] \\ &= pq^K\left[\frac{1}{1-q}\right] = \frac{pq^K}{p} = q^K \qquad (1) \end{aligned}$$

$$\begin{aligned} \therefore \quad p\,(X \geq K + t/X \geq K) &= \frac{p\,[X \geq K + t \cap X \geq K]}{p\,(X \geq K)} \\ &= \frac{p\,(X \geq K + t)}{p\,(X \geq K)} \\ &= \frac{q^{K+t}}{q^K} \qquad \text{(Using equation (1))} \\ &= q^t = p(X \geq t) \end{aligned}$$

$\therefore \quad p\,(X \geq K + t/X \geq K) = p\,(X \geq t)$.

Property 8: If X is an exponential distribution, then $p(X > K + t/X > K) = p(X > t)$ for any $K, t > 0$

Given that $X \sim Expo(\lambda)$, so that its pdf is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Now

$$p(X > K) = \int_K^\infty f(x)\,dx = \int_K^\infty \lambda e^{-\lambda x}\,dx$$

$$\begin{aligned} &= \lambda\left(\frac{e^{-\lambda x}}{-\lambda}\right)_K^\infty \\ p(X > K) &= e^{-K\lambda} \qquad (1) \end{aligned}$$

Now,

$$\begin{aligned} p(X > K + t/X > K) &= \frac{p[(X > K + t) \cap (X > K)]}{p(X > K)} \\ &= \frac{p(X > K + t)}{p(X > K)} \\ &= \frac{e^{-(K+t)\lambda}}{e^{-K\lambda}} \qquad \text{[using equation (1)]} \\ &= e^{-t\lambda} = p(X > t) \end{aligned}$$

Note: Properties (7) and (8) state the property of a geometric distribution and its continuous analog exponential distribution called *lack of memory*. In fact Property (7) states that the probability that atleast $K + t$ trials are required before the first success given that there have been K successive failures is equal to the unconditional probability that atleast t trials are needed before the first success. That is, the fact that though we have already observed K successive failures will not change the probability of the number of trials required to obtain the first success. Similarly, Property (8) states that if X represents the lifetime of a given component then an old functioning component has the same lifetime distribution as a new functioning component. This property is also called as *memory less property*. Indeed the converse of Property (7) and Property (8) are true which we prove now.

Property 9: If X is a non-negative integral valued random variable and $p(X \geq K + t/X \geq K) = p(X \geq t)$ then X is a geometric distribution.

We define $p(X = r) = p_r \quad r = 0, 1, 2, 3, \cdots$ and

$$q_k = p(X \geq k) = p_k + p_{k+1} + \cdots$$

Given that

$$\begin{aligned} p[X \geq k + t/X \geq k] &= p(X \geq t) \\ \Rightarrow \quad \frac{p[(X \geq k + t) \cap (X \geq k)]}{p(X \geq t)} &= p(X \geq t) \\ \Rightarrow \quad \frac{p(X \geq k + t)}{p(X \geq k)} &= p(X \geq t) \end{aligned}$$

In particular

$$\frac{p(X = k + t)}{p(X \geq k)} = p(X = t)$$

$$\Rightarrow \quad \frac{p_{k+t}}{q_k} = p_t$$

using our definition.

Hence $p_{k+1} = q_k p_t$ is true for all $t \geq 0$ and if $k \geq 0$. In particular $K = 1$, we have

$$\begin{aligned} p_{t+1} &= q_1 p_t \\ &= (p_1 + p_2 + \cdots)\, p_t \\ p_{t+1} &= (1 - p_0)\, p_t \end{aligned}$$

Hence $p_t = (1 - p_0)\, p_{t-1}$; $p_{t-1} = (1 - p_0)\, p_{t-2}$ so that

$$\begin{aligned} p_t &= (1 - p_0)\,(1 - p_0)\, p_{t-2} \\ &= (1 - p_0)^2\, p_{t-2} \end{aligned}$$

Similarly proceeding

$$\begin{aligned} p_t &= (1 - p_0)^t\, p_0 \\ \text{i.e.,}\quad p(X = t) &= p_0(1 - p_0)^t \quad t = 0, 1, 2, 3, \cdots \end{aligned}$$

which shows that X has a geometric distribution with parameter p_0.

Property 10: If a continuous RV $X > 0$ has memory less property then X has an exponential distribution.

Let X be the given RV with memory less Property. Let $f(x)$ and $F(x)$ be its pdf and cdf respectively. Given the space of X is $(0, \infty)$.

Consider

$$\begin{aligned} p\,(X > x + h/X > x) &= p\,(X > h) \\ &= \frac{p\,(X > x + h \cap X > x)}{p\,(X > x)} = p\,(X > h) \\ \Rightarrow &\quad p\,(X > x + h) = p\,(X > x)\, p\,(X > h) \\ \Rightarrow &\quad 1 - F\,(x + h) = [1 - F(x)]\,[1 - F(h)] \\ &\quad \text{Since, } F(x) = p(x \leq x) \\ \Rightarrow &\quad F\,(x + h) = F(x) + F(h) - F(x)F(h) \\ \Rightarrow &\quad \frac{F(x + h) - F(x)}{h} = \frac{F(h)\,[1 - F(x)]}{h} \end{aligned}$$

Letting $h \to \infty$, we have

$$\begin{aligned} F'(x) &= \lambda\,[1 - F(x)] \quad \text{where} \quad \lambda = \lim_{\lambda \to \infty} \frac{F(h)}{h} \\ \text{or} \qquad -\frac{F'(x)}{1 - F(x)} &= -\lambda \end{aligned}$$

Integrating we get, $1 - F(x) = Ke^{-\lambda x}$ but $F(0) = p(X \leq 0) = 0$ hence $1 - F(0) = K(1)$ so that $K = 1$.

$$\begin{aligned} \text{Hence}\quad 1 - F(x) &= e^{-\lambda x} \\ F(x) &= 1 - e^{-\lambda x} \end{aligned}$$

Differentiating, we get $\quad F'(x) = f(x) = \lambda e^{-\lambda x}\ x > 0, \lambda > 0.$

Hence, $X \sim Expo(\lambda)$.

At the end of this chapter we have tabulated the various aspects of discrete distributions (Table (1)) and continuous distributions (Table (2)) for a reference.

4.9 PROBLEMS

■ Example 4.31

If the MGF of a random variable X is $M_X(t) = \frac{1}{81}\left(2+e^t\right)^4$ find (a) $p(X=0)$, (b) mean and variance of X.

$$\text{Given,} \qquad M_X(t) = \frac{1}{81}\left(2+e^t\right)^4 = \frac{1}{3^4}\left(2+e^t\right)^4 = \left(\frac{2+e^t}{3}\right)^4 = \left(\frac{2}{3}+\frac{1}{3}e^t\right)^4$$

Hence, $X \sim b\left(4, \frac{1}{3}\right)$ so that the pmf of X is

$$f(x) = \begin{cases} \binom{4}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{4-x} & x = 0, 1, 2, 3, 4 \\ 0 & \text{elsewhere} \end{cases}$$

$$\therefore \quad p(X=0) = \left(\frac{2}{3}\right)^4 = \frac{16}{81} = 0.1975$$

$$\text{Mean} = np = 4 \times \frac{1}{3} = \frac{4}{3}$$

$$\text{Variance} = npq = 4 \times \frac{1}{3} \times \frac{2}{3} = \frac{8}{9}$$

■

■ Example 4.32

The MGF of a random variable X is $e^{-4(1-e^t)}$. Find $p(u - 2\sigma < X < u + 2\sigma)$.

Given that the MGF of X is

$$M_X(t) = e^{-4(1-e^t)} = e^{4(e^t-1)}$$

which is the MGF of a Poisson variable with $\lambda = 4$.

$\therefore \quad X \sim P(4)$ so that its pdf is

$$f(x) = \begin{cases} e^{-4}\frac{4^x}{x!} & x = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Also mean $(\mu) = \lambda = 4$ and variance $(\sigma^2) = 4$. Hence $\sigma = 2$.

Now, $$p(\mu - 2\sigma < X < \mu + 2\sigma) = p\left[4 - 2(2) < X < 4 + 2(2)\right]$$
$$= p(0 < X < 8) = \sum_{x=1}^{7} f(x)$$
$$= e^{-4}\left[4 + \frac{4^2}{2!} + \frac{4^3}{3!} + \cdots + \frac{4^7}{7!}\right]$$
$$= e^{-4}[4 + 8 + 10.6667 + 10.6667 + 8.5333$$
$$+5.6889 + 3.2508] = e^{-4}(50.8064) = 0.9306.$$

■

■ Example 4.33

If e^{3t+8t^2} is the MGF of a random variable X, find $p(-1 < X < 9)$.

$$\begin{aligned} M_X(t) &= e^{3t+8t^2} \\ &= e^{3t+\frac{16t^2}{2}} \end{aligned}$$

which is the MGF of a normal distribution with parameters $\mu = 3$ and $\sigma^2 = 16$ that is $X \sim n(3, 16)$.

$$\begin{aligned} \text{Now,} \qquad p(-1 < X < 9) &= p\left(\frac{-1-3}{4} < \frac{X-3}{4} < \frac{9-3}{4}\right) \\ &= p(-1 < Z < 1.5) \\ &= p(0 < Z < 1.5) + p(0 < Z < 1) \\ &= 0.4332 + 0.4713 \quad \text{From Table 2 of Appendix} \\ &= 0.9045. \end{aligned}$$

■

■ Example 4.34

If X and Y are two independent binomial variables with parameters $n_1 = 4$, $p = \frac{1}{3}$ and $n_2 = 6$, $p = \frac{1}{3}$ respectively. Compute (a) $p(X + Y = 3)$; (b) $p(X + Y \geq 2)$

Given that $X \sim b\left(4, \frac{1}{3}\right)$ and $Y \sim b\left(6, \frac{1}{3}\right)$. Hence, by the additive property $X + Y \sim b\left(10, \frac{1}{3}\right)$ so that the pmf of $Z = X + Y$ is

$$f_Z(Z) = \begin{cases} \binom{10}{Z}\left(\frac{1}{2}\right)^2\left(\frac{2}{3}\right)^{10-z} & Z = 0, 1, 2, \cdots, 10 \\ 0 & \text{elsewhere} \end{cases}$$

$$\begin{aligned} \text{(a)} \quad p(X + Y = 3) &= p(Z = 3) \\ &= \binom{10}{3}\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^{10-3} \\ &= 0.2601. \end{aligned}$$

$$\begin{aligned} \text{(b)} \quad p(X + Y \geq 2) &= p(Z \geq 2) = 1 - p(Z < 2) \\ &= 1 - [p(Z = 0) + p(Z = 1)] \\ &= 1 - \left[\left(\frac{2}{3}\right)^{10} + 10\left(\frac{1}{3}\right)\left(\frac{2}{3}\right)^9\right] \\ &= 0.8960. \end{aligned}$$

■

■ Example 4.35

Prove that the difference of two independent Poisson variables is not a Poisson variable.

Let $X \sim P(\lambda_1)$ and $Y \sim P(\lambda_2)$, so that $M_X(t) = e^{\lambda_1(e^t-1)}$ and $M_Y(t) = e^{\lambda_2(e^t-1)}$

$$\begin{aligned} \text{Now,} \quad M_{X-Y}(t) &= M_{X+(-Y)}(t) = M_X(t)\cdot M_{-Y}(t) \\ &\qquad \text{(Since } X \text{ and } Y \text{ are independent)} \\ M_X(t)\cdot M_Y(-t) &= e^{\lambda_1}(e^t-1)\cdot e^{\lambda_2(e^{-t}-1)} \\ &= e^{[\lambda_1(e^t-1)+\lambda_2(e^{-t}-1)]} \end{aligned}$$

which cannot be expressed in the form of the MGF of a Poisson distribution $\left[e^{\lambda(e^t-1)}\right]$. Hence $X - Y$ is not a Poisson variable. ■

■ Example 4.36

Prove that the difference of two independent normal vaiables is also a normal variable.

Let $X \sim n\left(\mu_1, \sigma_1^2\right)$, $Y \sim n\left(\mu_2, \sigma_2^2\right)$ be two independent normal variables.

$\therefore$ MGF of X and Y are

$$M_X(t) = e^{\mu_1 t+\sigma_1^2 t^2/2} \quad \text{and} \quad M_Y(t) = e^{\mu_2 t+\sigma_2^2 t^2/2}$$

Now, $M_{X-Y}(t) = M_X(t)\cdot M_{-Y}(t)$. Since X and Y are independent.

$$\begin{aligned} &= M_X(t)\cdot M_Y(-t) \\ &= \left[e^{\mu_1 t+\sigma_1^2 t^2/2}\right]\cdot\left[e^{\mu_2(-t)+\sigma_2^2 t^2/2}\right] \\ &= \left[e^{(\mu_1-\mu_2)t}+\frac{t^2}{2}\left(\sigma_1^2+\sigma_2^2\right)\right] \end{aligned}$$

which is of the form $e^{\mu t+\frac{\sigma^2 t^2}{2}}$, the MGF of the normal distribution.

$\therefore \quad X - Y$ is a normal distribution with parameters $\mu_1 - \mu_2$ and $\sigma_1^2 + \sigma_2^2$.

> **Remark**
>
> Examples 4.35 and 4.36 provide a difference between Poisson and normal distribution. In fact, linear combination of n independent normal variables is a normal variable (Refer Exercise 16).

■ Example 4.37

If X and Y are two independent geometric random variable with $p = \frac{1}{3}$, find (a) the distribution of $X + Y$and (b) the conditional distribution of X given $X + Y = 3$ [Refer property (6)]

Given that $X \sim Ge\left(\frac{1}{3}\right)$ and $Y \sim Ge\left(\frac{1}{3}\right)$

$\therefore q = 1 - p = \frac{2}{3}$ and pmf of X and Y are,

$$p(X = r) = p(Y = r) = \begin{cases} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right) & r = 0, 1, 2, 3, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Now, (a) $p(X + Y = K) = \sum_{r=0}^{K} p(X = r \cap Y = K - r)$

$$= \sum_{r=0}^{K} p(X = r) \cdot p(Y = K - r)$$

(Since X and Y are independent)

$$= \sum_{r=0}^{K} \left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right) \cdot \left(\frac{2}{3}\right)^{K-r} \left(\frac{1}{3}\right) \quad (\because \text{ Summation has } k+1 \text{ terms})$$

$$= \sum_{r=0}^{K} \left(\frac{2}{3}\right)^K \left(\frac{1}{3}\right)^2 = \left(\frac{2}{3}\right)^K \left(\frac{1}{3}\right)^2 (K+1)$$

(b) $p(X/X + Y = 3) = \dfrac{p(X = r \cap X + Y = 3)}{p(X + Y = 3)}$

$$= \frac{p(X = r) \cdot p(Y = 3 - r)}{p(X + Y = 3)}$$

$$= \frac{\left(\frac{2}{3}\right)^r \left(\frac{1}{3}\right) \cdot \left(\frac{2}{3}\right)^{3-r} \left(\frac{1}{3}\right)}{\left(\frac{2}{3}\right)^3 \left(\frac{1}{3}\right)^2 4} \quad \text{by (a)}$$

$$= \frac{1}{4}$$

■

■ Example 4.38

If $X \sim Expo(\lambda)$ with $p(X \leq 1) = p(X > 1)$. Find the mean and variance of X.

Since $X \sim Expo(\lambda)$, we have the pdf of X as

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Given that $p(X \leq 1) = p(x > 1)$

$$\Rightarrow \int_0^1 f(x)\, dx = \int_1^\infty f(x)\, dx$$

$$\Rightarrow \int_0^1 \lambda\, e^{-\lambda x} dx = \int_1^\infty e^{-\lambda x}\, dx$$

$$\Rightarrow \quad \left(-e^{-\lambda x}\right)_0^1 = \left(-e^{-\lambda x}\right)_1^\infty$$
$$\Rightarrow \quad (1 - e^{-\lambda}) = e^{-\lambda}$$
$$\therefore \quad e^{-\lambda} = \frac{1}{2}$$
$$\lambda = 0.693.$$

Hence, mean of $X = \frac{1}{\lambda} = 1.44$ and variance of $X = \frac{1}{\lambda^2} = 2.08$. ■

■ Example 4.39

Find the skewness and Kurtosis of the normal distribution.

The MGF of the normal distribution is

$$M_X(t) = e^{\mu t + \left(\frac{1}{2}\right)\sigma^2 t^2}$$

Expanding this as an exponential series, we have

$$\begin{aligned} M_X(t) &= 1 + \left(\mu t + \frac{1}{2}\sigma^2 t^2\right) + \frac{1}{2!}\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)^2 + \frac{1}{3!}\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)^3 \\ &\quad + \frac{1}{4!}\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)^4 + \cdots \\ &= 1 + \frac{t}{1!}(\mu) + \frac{t^2}{2!}\left(\mu^2 + \sigma^2\right) + \frac{t^3}{3!}\left(\mu^3 + 3\mu\sigma^2\right) + \\ &\quad \frac{t^4}{4!}\left(\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4\right) + \cdots \end{aligned}$$

Hence,

$$\begin{aligned} \mu_1' &= \text{coefficient of } \frac{t}{1!} \text{ in } M_X(t) = \mu \\ \mu_2' &= \text{coefficient of } \frac{t^2}{2!} \text{ in } M_X(t) \\ \mu_3' &= \text{coefficient of } \frac{t^3}{3!} \text{ in } M_X(t) = \mu^3 + 3\sigma^2\mu \\ \mu_4' &= \text{coefficient of } \frac{t^4}{4!} \text{ in } + 6M_X(t) \end{aligned}$$

$$\begin{aligned} \therefore \quad \mu_3 &= \mu_3' - 3\mu_2'\mu + 2\mu^3 \\ &= \mu^3 + 3\sigma^2\mu - 3\left(\mu^2 + \sigma^2\right)\mu + 2\mu^3 = 0 \\ \mu_4 &= \mu_4' - 4\mu_3'\mu + 6\mu_2'\mu^2 - 3\mu^4 \\ &= \left(\mu^4 + 6\mu^2\sigma^2 + 3\sigma^4\right) - 4\left(\mu^3 + 3\sigma^2\mu\right)\mu + 6\left(\mu^2 + \sigma^2\right)\mu^2 - 3\mu^4 \\ &= 3\sigma^4 \end{aligned}$$

Therefore, the coefficient of skewness $= \dfrac{\mu_3^2}{\mu_2^3} = 0$

and the coefficient of kurtosis $= \dfrac{\mu_4}{\mu_2^2} = \dfrac{3\sigma^4}{\sigma^4} = 3$. ■

■ Example 4.40

If $X \sim b(n, p)$ show that the MGF of$Z = \frac{X-np}{\sqrt{npq}}$ approaches the MGF of $n(0, 1)$ when $n \to \infty$.

$$\text{Since,} \qquad X \sim b(n, p) \text{ its MGF is } M_X(t) = \left(q + pe^t\right)^n \tag{1}$$

$$\begin{aligned}
\text{Now,}\ M_Z(t) &= E\left[e^{tz}\right] \\
&= E\left[e^{t\left(\frac{X-np}{\sqrt{npq}}\right)}\right] = E\left[e^{\frac{tX}{\sqrt{npq}}} \cdot e^{-\frac{tnp}{\sqrt{npq}}}\right] \\
&= e^{\frac{-tnp}{\sqrt{npq}}} e\left[e^{\frac{tX}{\sqrt{npq}}}\right] \\
&= e^{\frac{-tnp}{\sqrt{npq}}} e\left[e^{\frac{t}{\sqrt{npq}}X}\right] = e^{\frac{-tnp}{\sqrt{npq}}} M_X\left(\frac{t}{\sqrt{npq}}\right) \\
&= e^{-\frac{tnp}{\sqrt{npq}}}\left(q + pe^{\frac{t}{\sqrt{npq}}}\right)^n \cdot \quad \text{[using equation (1)]} \\
&= \left(e^{-\frac{tp}{\sqrt{npq}}}\right)^n \left(q + pe^{\frac{t}{\sqrt{npq}}}\right)^n \\
&= \left[qe^{-\frac{tp}{\sqrt{npq}}} + pe^{-\frac{tp}{\sqrt{npq}}} \cdot e^{\frac{t}{\sqrt{npq}}}\right]^n \\
&= \left[qe^{-\frac{tp}{\sqrt{npq}}} + pe^{\frac{t}{\sqrt{npq}}(1-p)}\right]^n \\
&= \left[qe^{-\frac{tp}{\sqrt{npq}}} + pe^{\frac{t}{npq}q}\right]^n \\
&= \left\{q\left[1 - \frac{tp}{\sqrt{npq}} + \frac{1}{2!}\left(\frac{tp}{\sqrt{npq}}\right)^2 - \frac{1}{3!}\left(\frac{tp}{\sqrt{npq}}\right)^3 + \frac{1}{4!}\left(\frac{tp}{\sqrt{npq}}\right)^4 - \cdots\right]\right. \\
&\qquad \left. + p\left[1 + \frac{tq}{\sqrt{npq}} + \frac{1}{2!}\left(\frac{tq}{\sqrt{npq}}\right)^2 + \frac{1}{3!}\left(\frac{tq}{\sqrt{npq}}\right)^3 + \frac{1}{4!}\left(\frac{tq}{\sqrt{npq}}\right)^4 + \cdots\right]\right\}^n \\
&= \left[1 + \frac{t^2}{2!}\left(\frac{pq^2 + p^2q}{npq}\right) + \frac{t^3}{3!}\left(\frac{pq^3 - p^3q}{npq\sqrt{npq}}\right) + \frac{t^4}{4!}\left(\frac{pq^4 + p^4q}{(npq)^2}\right) + \cdots\right]^n \\
&= \left[1 + \frac{t^2/2!}{n}(p+q) + \frac{t^4/3!}{n}\left(\frac{q^2 - p^2}{\sqrt{npq}}\right) + \frac{t^2/4!}{n}\left(\frac{q^3 + p^3}{npq}\right) + \cdots\right]^n \\
&= \left[1 + \frac{t^2/2}{n} + \frac{1}{n}(\phi(n))\right]^n
\end{aligned}$$

$$\begin{aligned}
\text{where,} \qquad \phi(n) &= \frac{t^3/3!\left(q^2 - p^2\right)}{\sqrt{npq}} + \frac{t^4/4!\left(q^3 - p^3\right)}{\sqrt{npq}} + \cdots \\
\text{and} \qquad \lim_{n\to\infty} \phi(n) &= 0
\end{aligned}$$

$$\text{Hence,} \qquad \lim_{n\to\infty} M_Z(t) = \lim_{n\to\infty}\left[1 + \frac{t^2/2}{n} + \frac{1}{n}(\phi(n))\right]^n = e^{t^2/2}$$

which is the MGF of the standard normal distribution. ■

■ Example 4.41

If $X \sim P(\lambda)$, then show that MGF of $Z = \frac{X-\lambda}{\sqrt{\lambda}}$ approaches the MGF of $n(0, 1)$ when $\lambda \to \infty$.

Since $X \sim P(\lambda)$ the MGF of X is

$$M_X(t) = e^{\lambda(e^t-1)} \tag{1}$$

Consider

$$\begin{aligned}
M_Z(t) &= E[e^{tZ}] \\
&= \left[e^{t\left(\frac{X-\lambda}{\sqrt{\lambda}}\right)}\right] = E\left[e^{tX/\sqrt{\lambda}}\, e^{-t\lambda/\sqrt{\lambda}}\right] \\
&= e^{-t\sqrt{\lambda}} e^{\lambda\left(e^{t\sqrt{\lambda}}-1\right)} \\
&= e^{-t\sqrt{\lambda}} E\left[e^{\frac{t}{\sqrt{\lambda}}X}\right] \\
&= e^{-t\sqrt{\lambda}} M_X\left(\frac{t}{\sqrt{\lambda}}\right) \\
&= e^{-t\sqrt{\lambda}} e^{\left(\lambda e^{t\sqrt{\lambda}}-\lambda\right)} \\
&= e^{\left(-t\sqrt{\lambda}+\lambda e^{\frac{t}{\sqrt{\lambda}}}-\lambda\right)} \qquad (2)
\end{aligned}$$

Now $-t\sqrt{\lambda} + \lambda e^{t\sqrt{\lambda}} - \lambda$

$$\begin{aligned}
&= -t\sqrt{\lambda} + \lambda\left[1 + \left(\frac{t}{\sqrt{\lambda}}\right) + \frac{1}{2!}\left(\frac{t}{\sqrt{\lambda}}\right)^2 + \frac{1}{3!}\left(\frac{t}{\sqrt{\lambda}}\right)^3 + \right. \\
&\qquad \left. \frac{1}{4!}\left(\frac{t}{\sqrt{\lambda}}\right)^4 + \cdots\right] - \lambda \\
&= -t\sqrt{\lambda} + \lambda + t\sqrt{\lambda} + t^2/2 + \frac{t^3/3!}{\sqrt{\lambda}} + \frac{t^4/4!}{\lambda} + \cdots - \lambda \\
&= \frac{t^2}{2} + \frac{t^3}{3!}.\frac{1}{\sqrt{\lambda}} + \frac{t^4}{4!}.\frac{1}{\lambda} + \cdots \\
&= \frac{t^2}{2} \quad \text{when} \quad \lambda \to \infty
\end{aligned}$$

$\therefore$ from equation (2), we have

$$\lim_{\lambda\to\infty} M_Z(t) = e^{t2/2}$$

which is the MGF of the standard normal distribution. ■

Remark

Examples 4.40 and 4.41 are infact illustrate the uniqueness property of MGF and more over they show the use of the normal distribution. These results apply when $n \to \infty$, but the normal distribution is often used to approximate binomial and Poisson probabilities even when $n > 30$. We will discuss this idea more in our following chapters, of course we prove this results still more in an easier (and of course more natural) method using one of the most important and interesting theorems called *central limit theorem*.

4.10 MULTINOMIAL AND BIVARIABLE NORMAL

We conclude this discussion with two distributions namely multinomial distribution and bivariate normal distribution. The names suggest our idea. An immediate generalization of binomial distribution arises when each trial has more than two possible outcomes, the probabilities of the respective outcomes are the same for each trial, and the trials are independent. For example a group of students interviewed by an opinion poll to find whether a new method of examination is rated excellent, above average, average or inferior, then each of these four mutually exclusive outcomes have same probability say p_1, p_2, p_3, p_4 with $p_1 + p_2 + p_3 + p_4 = 1$. If n students are interviewed and if we are interested in finding the probability of x_1 students favour 'excellent' x_2 students favour 'above average' x_3 students favour 'average' and x_4 students favour 'inferior' (here $x_1 + x_2 + x_3 + x_4 = n$), then $p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \cdot p_4^{x_4}$ is the required value (Recall the discussion in binomial distribution). Further, the number of ways of getting x_1 students favour 'excellent', x_2 students favour 'average' and so on is

$$\frac{n!}{x_1! \cdot x_2! \cdot x_3! \cdot x_4!}$$

Hence, the required probability is

$$\frac{n!}{x_1! \cdot x_2! \cdot x_3! \cdot x_4!}\, p_1^{x_1} \cdot p_2^{x_2} \cdot p_3^{x_3} \cdot p_4^{x_4}$$

where $p_1 + p_2 + p_3 + p_4 = 1$ and $x_1 + x_2 + x_3 + x_4 = n$. Generalising this example we define the following.

Definition

The random variables $X_1, X_2, \cdots, X_n$ is said to have a **multinomial distribution** if their joint probability distribution is given by

$$f(x_1, x_2 \cdots x_K;\, p_1 p_2 \cdots p_K) = \frac{n!}{x_1! x_2! \cdots x_K!} p_1^{x_1} \cdot p_2^{x_2} \cdots p_k^{x_k}$$

for $x_i = 0, 1, 2, \cdots, n$ for each i and $x_1 + x_2 + \cdots + x_K = n$ and $p_1 + p_2 + \cdots p_K = 1$

In the definition of multinomial distribution we refer to the outcomes as being of the first kind, the second kind and the K^{th} kind. The corresponding probability is getting x_1 outcomes of the first kind, x_2 outcomes of the second kind $\cdots$ and x_K outcomes of the K^{th} kind.

■ Example 4.42

In a certain survey, channel 9 intends to study the audience preference, for three of its programmes during prime time. It observes that 50% viewing audience, 30% viewing audience and 20% viewing audience prefer programme I, II and III respectively. If 10 viewers are randomly selected, find the probability that during prime time 5 will be watch programme I, two will watch II and 3 will be watching III.

Here $$n = 10(5+2+3)$$

$p_1 = p(\text{viewers favour programme } I) = 0.50$, similarly $p_2 = 0.30$ and $p_3 = 0.20$. Also $x_1 = 5$, $x_2 = 2$, $x_3 = 3$. Hence, the required probability is

$$\begin{aligned} f\,(5, 2, 3; 0.5, 0.3, 0.2) &= \frac{10!}{5!\,2!\,3!}\,(0.5)^5\,(0.3)^2(0.2)^3 \\ &= 0.0567 \end{aligned}$$

MGF of multinomial distribution

$$\begin{aligned} M_x(t) &= E\left(e^{tx}\right) \\ &= E\left[e^{\sum\limits_{t=1}^{k} t_i x_i}\right] \\ &= \sum_x \left[\frac{n!}{x_1!\,x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} e^{\left(\sum\limits_{i=1}^{k} t_i x_i\right)}\right] \\ &= \sum_x \left[\frac{n!}{x_1!\,x_2!\cdots x_k!} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} e^{t_1 x_1} \cdot e^{t_2 x_2} \cdots e^{t_k x_k}\right] \\ &= \sum_x \left[\frac{n!}{x_1!\,x_2!\cdots x_k!} \left(p_1 e^{t_1}\right)^{x_1} \left(p_1 e^{t_2}\right)^{x_2} \cdots \left(p_k e^{t_k}\right)^{x_k}\right] \\ &= \left(p_1 e^{t_1} + p_2 e^{t_2} + \cdots + p_k e^{t_k}\right)^n \cdot x = (x_1, x_2, \cdots x_k) \end{aligned}$$

From this we can find the MGF of X_i using

$$\begin{aligned} M_{X_i}\,(t_i) &= M_X\,(0, 0\cdots 0, \underset{i^{\text{th}}\text{ place }\uparrow}{1}, 0\cdots 0) \\ &= p_1 + p_2 + \cdots + p_{i-1} + p_i e^{t_i} + p_{i+1} + \cdots + p_k \\ &= \left(q_i + p_i e^{t_i}\right)^n \; q_i = p_1 + p_2 + \cdots + p_{i-1} + p_{i+1} + \cdots + p_k \\ &= 1 - p_i \end{aligned}$$

which is the MGF of a binomial distribution. Hence $X_i \sim (n, pi)$ so that the marginal probability distribution of X_i is

$$f_i\,(X_i) = \begin{cases} \begin{pmatrix} n \\ x \end{pmatrix} p_i^x q_i^{n-x} & x = 0, 1, 2, \cdots, n \\ 0 & \text{elsewhere} \end{cases}$$

and mean of $X_i = np_i$ variance of $X_i = np_i q_i$, $i = 1, 2, 3, \cdots, k$.

Further, we can find the Cov $\left(X_i, X_j\right)$ $i \neq j$ and $i, j = 1, 2, \cdots, k$ using the joint MGF of multinomial distribution, using the fact that

$$
\begin{aligned}
E\left(X_i X_j\right) &= \left[\frac{\partial^2}{\partial t_i \partial t_j}\left(M_X(t)\right)\right]_{t=0} i \neq j \\
&= \left[\frac{\partial^2}{\partial t_i \partial t_j}\left(p_1 e^{t_1} + p_2 e^{t_2} + \cdots + p_k e^{t_k}\right)\right]_{t=0} \\
&= \left[n(n-1)\, p_i p_j \left(p_1 e^{t_1} + p_2 e^{t_2} + \cdots + p_k e^{t_k}\right)^{n-2}\right]_{t=0} \\
&= n(n-1)\, p_i q_j \\
\therefore \quad \operatorname{Cov}\left(X_i X_j\right) &= E\left(X_i X_j\right) - E\left(X_i\right) E\left(X_j\right) \\
&= n(n-1)\, p_i p_j - \left(n p_i\right)\left(n p_j\right) \\
&= n^2 p_i p_j - n p_i p_j - n^2 p_i p_j = -n p_i p_j \\
\text{Hence,} \qquad \rho_{X_i X_j} &= \frac{-n p_i p_j}{\sqrt{n p_i q_j}\sqrt{n p_j q_j}} \\
&= -\sqrt{\frac{p_i p_j}{q_i q_j}}
\end{aligned}
$$

Now, we shall attempt to generalise the normal distribution but limit ourselves to the two variable case. The reason is we need to present the multivariate normal distribution essentially in a matrix form, which we do not try in our discussion.

Definition

Two random variables X and Y have a **bivariate normal distribution**, if their joint probability density is given by

$$
\begin{aligned}
f(x, y) = \; & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-p^2}} Exp{-1/2}\left(1-p^2\right)\left[\left(\frac{x-\mu_1}{\sigma_1^2}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right. \\
& \left. -2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right]
\end{aligned}
$$

for $-\infty < x < \infty$ and $-\infty < y < \infty$ and $\sigma_1^2 = V(X)$, $\sigma_2^2 = V(Y)$ and $\mu_1 = E(X)$, $\mu_2 = E(y)$ and ρ is the correlation coefficient of X and Y.

The bivariate normal distribution has many important properties. However, we shall study about the marginal densities of X and Y and the conditional densities.

Result 1: If (X, Y) is a bivariate normal distribution then the marginal density of X is a normal distribution with mean μ_1 and variance σ_1^2, we have

$$
\begin{aligned}
f(x, y) = \; & \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} Exp{-1/2}\left(1-\rho^2\right)\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right. \\
& \left. -2p\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right] \quad -\infty < x < \infty \quad \text{and} \quad -\infty < y < \infty
\end{aligned}
$$

If $f_X(x)$ is the marginal density of X, then

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\, dy$$

$$= \frac{1}{\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right)^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right]} dy \qquad (1)$$

Applying the method of completing squares, we have

$$\left(\frac{Y-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{X-\mu_1}{\sigma_1}\right)\left(\frac{Y-\mu_2}{\sigma_2}\right) = \left[\left(\frac{Y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{X-\mu_1}{\sigma_1}\right)\right]^2 - \rho^2\left(\frac{X-\mu_1}{\sigma_1}\right)^2$$

Hence from equation (1),

$$f_X(x) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{X-\mu_1}{\sigma_1}\right)^2} e^{\frac{\rho^2}{2(1-\rho^2)}\left(\frac{X-\mu_1}{\sigma_1}\right)^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{X-\mu_1}{\sigma_1}\right)\right]^2} dy$$

$$= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left[\frac{x-\mu_1}{\sigma_1}\right]^2} \int_{-\infty}^{\infty} e^{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{Y-\mu_2}{\sigma_2}\right) - \rho\left(\frac{X-\mu_1}{\sigma_1}\right)\right]^2} dy$$

If we substitute $Z = \frac{Y-\mu_2}{\sigma_2}$ so that $dy = \sigma_2\, dz$ and $W = \frac{X_1-\mu_1}{\sigma_1}$, then the integral becomes

$$f_X(x) = \frac{1}{2\pi\sigma_1\sqrt{1-\rho^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\left(1-\rho^2\right)}(Z-\rho W)^2} dz$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{(Z-\rho W)^2}{2\left(1-\rho^2\right)}}\, dz \qquad (3)$$

The integral in equation (3) can be identified as the integral of a normal distribution and hence it is equal to 1.

We get, $$f_X(x) = \frac{1}{\sigma_1\sqrt{2\pi}}\, e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}} \quad -\infty < X < \infty$$

which is the pdf of $n\left(\mu_1, \sigma_1^2\right)$ and hence $X \sim n\left(\mu_1, \sigma_1^2\right)$. Similarly we can prove that $Y \sim n\left(\mu_2, \sigma_2^2\right)$.

We Result 2 derive the conditional density of X given Y using the fact that

$$
\begin{aligned}
f_{X/Y}(x/y) &= \frac{f(x, y)}{f_Y(y)} \\
& \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}}\left[W^2 - 2\rho W Z + Z^2\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{1}{2}Z^2} \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\, e^{-\frac{1}{2(1-\rho^2)}}\left[W^2 - 2\rho W Z + Z^2 - \left(1-\rho^2\right)Z^2\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\, e^{-\frac{1}{2\left(1-\rho^2\right)}}\left[W^2 - 2\rho W Z + \rho^2 Z^2\right] \\
&= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\, e^{-\frac{1}{2\left(1-\rho^2\right)}}\left[W - \rho_Z\right]^2 \qquad (4)
\end{aligned}
$$

But
$$
\begin{aligned}
W - \rho Z &= \left(\frac{X-\mu_1}{\sigma_1}\right) - \rho\left(\frac{Y-\mu_2}{\sigma_2}\right) \\
&= \frac{1}{\sigma_1}\left[X - \left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho\,(Y-\mu_2)\right)\right]
\end{aligned}
$$

Hence from equation (1), we have

$$
\begin{aligned}
f_{X/Y}(x/y) &= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\, e^{-\frac{\left[X-\left(\mu_1+\rho\frac{\sigma_1}{\sigma_2}(Y-\mu_2)\right)\right]^2}{2\sigma_1^2\left(1-\rho^2\right)}} \\
&= \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{(X-\mu^2)}{2\sigma^2}}
\end{aligned}
$$

where $\mu = \mu_1 + \rho\dfrac{\sigma_1}{\sigma_2}(Y-\mu_2)$ and $\sigma = \sigma_1\sqrt{1-\rho^2}$

so that it can be observed that $f_{X/Y}(x/y)$ is a normal distribution with mean μ and standard deviation σ as we defined in the above equations. The corresponding result for the conditional density of Y given X is obtained in a similar method. That is the conditional density of Y given X is a normal density with parameters mean $M = \mu_2 + \rho\frac{\sigma_2}{\sigma_2}(x-\mu_1)$ and variance $\sigma^2 = \sigma_2^2\left(1-\rho^2\right)$

(or) $$f_{Y/X}(y/x) = \frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}} e^{-\frac{\left[Y-\left(\mu_2+\rho\frac{\sigma_2}{\sigma_1}(x-\mu_1)\right)\right]^2}{\sigma_2^2\sqrt{1-\rho^2}}}$$

Also, it can easily be observed that the above discussion provides the conditional mean and variance of the bivariate normal distribution X and Y.

Before we conclude let us consider a special case of the bivariate normal distribution. That is when $\rho = 0$ and $\sigma_1 = \sigma_2 = \sigma$ which is called as a *circular normal* distribution with density

$$f(x, y) = \frac{1}{2\pi\sigma_2} e^{-\frac{1}{2\sigma^2}}\left[(X-\mu_1)^2 + \left(Y-\mu_2\right)^2\right]$$

4.11 Additional Examples

■ Example 4.43

If the probability that a bomb will hit a target is 0.8, what is the probability that out of 12 bombings exactly 3 are missed?

Let X be the number of misses. Hence, this is a binomial variable with $n = 12$, $p = 0.2$, that is $X \sim b(12, 0.2)$. Hence,

$$f(x) = \begin{cases} \begin{pmatrix} 12 \\ x \end{pmatrix} (0.2)^x (0.8)^{12-x} & x = 0, 1, 2, \cdots, 12 \\ 0 & \text{elsewhere} \end{cases}$$

$\therefore$ Required probability is $p(X = 3)$

$$\begin{aligned} &= \begin{pmatrix} 12 \\ 3 \end{pmatrix} (0.2)^3 (0.8)^{12-3} \\ &= 0.2362. \end{aligned}$$

■

■ Example 4.44

An item is produced with 2% defectives. A quality control inspector is examining the items by taking them at random. What is the probability that atleast 4 items are to be examined in order to get 2 defectives.

Let the success event be getting second defective item.
Hence, this is a negative binomial distribution with $r = 2$ and $p = 2\% = 0.02$

$$\therefore \quad f(x) = \begin{cases} \begin{pmatrix} x + 2 - 1 \\ 2 - 1 \end{pmatrix} (0.02)^2 (0.98)^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Here X is the number of non-defectives chosen or 'failure' event.

$$f(x) = \begin{cases} (x + 1)(0.02)^2 (0.98)^x & x = 0, 1, 2, \cdots \\ 0 & \text{elsewhere} \end{cases}$$

Required probability $=$ p(atleast 4 items are to be examined)

$$\begin{aligned} &= p(X \geq 2) \\ &= 1 - p(X < 2) \\ &= 1 - [p(X = 0) + p(X = 1)] \\ &= -[(0.02)^2 + 2(0.02)^2 0.98] \\ &= 0.9988. \end{aligned}$$

■

■ Example 4.45

A number of people are swimming across a lake. Each one has a probability 0.4 of successfully completing it. Compute the probability that the 10th person is (a) the first one to complete the swimming successfully, (b) the third one to complete the swimming successfully.

Let the success event be swimming across the lake successfully, so that $p = 0.4$.

(a) Here we need to calculate the first success in the tenth trial. Hene X has a geometric distribution with $p = 0.4$ and $q = 1 - p = 0.6$

$$\begin{aligned} \therefore \quad \text{Required probability} &= q^9 p \\ &= (0.6)^9(0.4) \\ &= 0.004. \end{aligned}$$

(b) This is a negative binomial distribution with $r = 3$ and $x + r = 10$ so that $x = 7$.

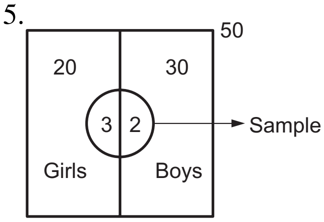
$$\begin{aligned} \therefore \quad \text{Required probability} &= \binom{7+3-1}{3-1}(0.4)^3(0.6)^7 \\ &= 0.0645. \end{aligned}$$

■

■ Example 4.46

In a class of 50 students 20 are girls and 30 are boys. If 5 students are selected at random what is the probability that 2 are boys?

This is a hypergeometric distribution with $N = 50$, $M = 30$, $n = 5$.



$$\therefore \quad f(x; N, M, n) = \frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$$

Here $x = 2$,

$$f(2) = \frac{\binom{30}{2}\binom{20}{3}}{\binom{50}{5}} = 0.2341$$

<u>Note:</u> Example 4.46 can be approximated with binomial distribution. Let X be the number of boys in the sample of 5, so that $n = 5$ and $p = \frac{3}{5}$. The required probability is $\binom{5}{2}\left(\frac{3}{5}\right)^2\left(\frac{2}{5}\right)^3 = 0.2304$.

■ Example 4.47

If the duration of a shower in an island is exponentially distributed with $\lambda = \frac{1}{5}$, (a) out of 3 showers, what is the probability that not more that 2 will last for 10 minutes or more? Also compute the probability that (b) a shower will last atleast 2 minutes more given that it has already lasted for 5 minutes (c) a shower will not last more than 6 minutes more if it has already lasted for 3 minutes.

Let X be the duration of a shower. Give that $X \sim Expo\left(\frac{1}{5}\right)$ hence the pdf of X is

$$f(x) = \begin{cases} \frac{1}{5}e^{-\frac{1}{5}x} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

(a) Here we count the number of successes if the success event is that the shower will last for 10 minutes or more.

$$\text{Hence,} \qquad p = p(X \geq 10) = \int_{10}^{\infty} \frac{1}{5}e^{-\frac{1}{5}x} = 0.1353.$$

Hene out of $n = 3$ showers, we need to compute the probability of getting atmost 2 successes. If Y is the number of successs then $Y \sim b(5, p)$ $p = 0.1353$ and $q = 0.8647$.

$$\begin{aligned} \therefore \quad \text{Required probability} &= \sum_{x=0}^{2} \binom{5}{x} (0.1353)^x (0.8647)^{5-x} \\ &= (0.8647)^5 + 5(0.1353)(0.8647)^4 + 10(0.1353)^2(0.8647)^3 \\ &= 0.9799. \end{aligned}$$

(b) $$p(X > 7/X > 5) = \frac{p(X > 7 \cap X > 5)}{p(X > 5)} = \frac{p(X > 7)}{p(X > 5)} = \frac{\int_7^{\infty} \frac{1}{5}e^{-\frac{1}{5}x}dx}{\int_5^{\infty} \frac{1}{5}e^{-\frac{1}{5}x}dx} = 0.6703.$$

(c) $$p(X < 9/X > 3) = \frac{p(3 < X < 9)}{p(X > 3)} = \frac{\int_3^{9} \frac{1}{5}e^{-\frac{1}{5}x}}{\int_3^{\infty} \frac{1}{5}e^{-\frac{1}{5}x}} = 0.6988.$$

***Note*:** In Example 4.47 (b), we can apply memory less property of exponential distribution and hence the required probability is $p(X > 2) = 0.6703$ i.e., $p[X > 5 + 2/X > 5]$.

■ Example 4.48

The demand for a certain item per day is distributed as a gamma distribution with parameters $\alpha = 2$ and $\beta = 4$ what is the probability that there will be a demand for atleast 10 units on a particular day?

Let X be the demand for the item so that $X \sim G(x; 2, 4)$

pdf of X, is

$$f(x) = \begin{cases} \frac{1}{4^2\lceil 2} x e^{-x/4} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$= \begin{cases} \frac{1}{16} x e^{-x/4} & x > 0 \\ 0 & \text{elsewhere} \end{cases}$$

$$p(X \geq 10) = \int_{10}^{\infty} \frac{1}{16} x\, e^{-x/4}\, dx$$

■ Example 4.49

Let X and Y have a bivariate normal distribution with parameters $\mu_1 = 3, \mu_2 = 1, \sigma_1^2 = 16, \sigma_2^2 = 25$ and $\rho = \frac{3}{5}$. Compute the following probabilities (a) $p(3 < Y < 8)$, (b) $p(3 < Y < 8/x = 7)$, (c) $p(-3 < X < 3)$, (d) $p(-3 < X < 3/Y = -4)$.

If X and Y have a bivariate normal distribution then, we have

(i) $X \sim n\left(\mu_1, \sigma_1^2\right)$

(ii) $Y \sim n\left(\mu_2, \sigma_2^2\right)$

(iii) $Y/X \sim n\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2\left(1 - \rho^2\right)\right]$

(iv) $X/Y \sim n\left[\mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(Y - \mu_2\right), \sigma_1^2\left(1 - \rho^2\right)\right]$

Hence, we compute probabilities in (a), (b), (c) and (d) using (ii), (iii), (i) and (iv) respectively.

(a) $Y \sim n(1, 25)$

$$\therefore \quad p(3 < Y < 8) = p\left(\frac{3-1}{5} < Z < \frac{8-1}{5}\right)$$
$$= p(0.4 < Z < 1.4) = 0.264$$

(Refer normal table to get the probabilities)

(b) $(Y/X = 7) \sim n\left[1 + \frac{3}{5} \cdot \frac{5}{4}(7-3), 25\left(1 - \frac{9}{25}\right)\right] = n(4, 16)$

$$\therefore \quad p\,(3 < Y < 8/X = 7) = p\left(\frac{3-4}{4} < Z < \frac{8-4}{4}\right)$$
$$= p\,(-0.25 < Z < 1) = 0.44.$$

(c) $X \sim n(3, 16)$

$$\therefore \quad p\,(-3 < X < 3) = p\left(\frac{-3-3}{4} < Z < \frac{3-3}{4}\right)$$
$$= p\,(-1.5 < Z < 0) = 0.433$$

(d) $(X/Y=-4) \sim n\left(3+\frac{3}{5}\frac{4}{5}(-4-1), 16\left(1-\frac{9}{25}\right)\right) = n\left(\frac{3}{5}, \frac{256}{25}\right)$

$$\begin{aligned}\therefore \quad p\,(-3 < X < 3/Y=-4) &= p\left(\frac{-3-3/5}{16/5} < Z < \frac{3-3/5}{16/5}\right)\\ &= p\,(-1.125 < Z < 0.75) = 0.644.\end{aligned}$$

■ Example 4.50

Let X and Y have a bivariate normal distribution with parameters $\mu_1 = 5$, $\mu_2 = 10$, $\sigma_1^2 = 1$, $\sigma_2^2 = 25$ and $\rho > 0$. If $p\,(4 < Y < 16/x = 5) = 0.954$ find ρ.

Since, $\quad (Y/X = x) \sim n\left[\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x-\mu_1), \sigma_2^2\left(1-\rho^2\right)\right]$

we have, $\quad (Y/X=5) \sim n\left[10+\rho\frac{5}{1}(5-5), 25\left(1-\rho^2\right)\right]$

that is, $\quad (Y/X=5) \sim n\left[10, 25\left(1-\rho^2\right)\right]$.

Given, $p(4 < y < 16/x = 5) = 0.954$

$$\Rightarrow \quad p\left(\frac{4-10}{5\sqrt{1-\rho^2}} < Z < \frac{16-10}{5\sqrt{1-\rho^2}}\right) = 0.954$$

$$\Rightarrow \quad p\left(\frac{-1.2}{\sqrt{1-\rho^2}} < Z < \frac{1.2}{\sqrt{1-\rho^2}}\right) = 0.954$$

$$\Rightarrow \quad 2p\,(0 < Z < Z_1) = 0.954 \quad \text{where} \quad Z_1 = \frac{1.2}{\sqrt{1-\rho^2}}$$

$$\Rightarrow \quad p\,(0 < Z < Z_1) = 0.477$$

Refer to the standard normal table, we have

$$\begin{aligned} Z_1 &= 2\\ \therefore \quad \frac{1.2}{\sqrt{1-\rho^2}} &= 2\end{aligned}$$

Solving this we get $\rho = \frac{4}{5}$.

■ Example 4.51

Find correlation coefficient between X and Y for the bivariate normal distribution.

$$f(x,y) = \frac{1}{2\pi\sqrt{3}}\,e^{-\left[\frac{(2x-y)^2+2xy}{6}\right]} \quad -\infty < x, y < \infty$$

The joint pdf of X and Y, is

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}}\,e^{-\frac{1}{2(1-\rho^2)}\left[u^2-2puv+V^2\right]}$$

$$\text{where} \qquad U = \frac{X - \mu_1}{\sigma_1} \quad \text{and} \quad V = \frac{y - \mu_2}{\sigma_2}$$

If the exponent of e of this pdf is expanded, then the coefficients of x, y and xy are respectively,

$$\frac{1}{2\sigma_1^2\left(1-\rho^2\right)}, \frac{1}{2\sigma_2^2\left(1-\rho^2\right)} \quad \text{and} \quad -\frac{\rho}{\left(1-\rho^2\right)\sigma_1\sigma_2}$$

(neglecting the '$-$' sign).

Now, consider the given pdf and the exponent of e is $\frac{4}{6}x^2 + \frac{1}{6}y^2 - \frac{2}{6}xy$. Hence, comparing the coefficients of x^2, y^2 and xy, we have

$$\frac{1}{2\sigma_1^2\left(1-\rho^2\right)} = \frac{4}{6} = \frac{2}{3} \tag{1}$$

$$\frac{1}{2\sigma_2^2\left(1-\rho^2\right)} = \frac{1}{6} \tag{2}$$

$$\frac{\rho}{\left(1-\rho^2\right)\sigma_1\sigma_2} = \frac{1}{3} \tag{3}$$

equation (1) $\div$ equation (2) $\Rightarrow$ $\frac{\sigma_2^2}{\sigma_1^2} = 4$ or $\sigma_2 = 2\sigma_1$

$\therefore$ Equation (3) becomes $\frac{\rho}{2(1-\rho^2)\sigma_1^2} = \frac{1}{3}$ using equation (1) in this equation, we have

$$\rho \cdot \frac{2}{3} = \frac{1}{3}$$

$$\therefore \qquad \rho = \frac{1}{2}$$

***Note*:** In Example 4.51, infact we can find σ_1 and σ_2 also, that is $\sigma_1 = 1$ and $\sigma_2 = 2$ (verify). Moreover, we can observe that $\mu_1 = \mu_2 = 0$.

Remark

The problems based on special distributions can be listed in two ways. One is to identify the role and utility of a particular distribution and the other is to apply or make use of their properties, and characteristics, such as mean variance ans so on. Hence, the Exercise in this chapter aims to test our ideas in this direction. However an aspect regarding this special distribution namely distribution function is included in the exercise set. Also, we do not maintain any order in listing the problems to match with the discussions we had in the chapter.

Table 4.2

	Binomial Distribution	Poisson Distribution	Negative Binomial Distribution	Geometric Distribution	Hyper Geometric Distribution
Description	Number of success (x)	Number of successes (x)	The number of trial in which rth	The number of trial in which First	Sampling without replacement
PDF and Range	$\binom{n}{x} p^x q^{n-x}$ $x = 0, 1, 2, n$	$e^{-\lambda} \frac{\lambda^x}{x!}$ $x = 0, 1, 2 \cdots$	$\binom{x+r-1}{r-1} p^r q^x$ $x = 0, 1, 2, 3 \cdots$	$q^x p$ $x = 0, 1, 2, \cdots$	$\frac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$ $x = 0, 1, 2 \cdots n$
Symbol and Parameter(S)	$x \sim b(n, p)$ n and p	$X \sim p(\lambda)$ $\lambda = np$	$X \sim nb(x; r, p)$ r and p	$X \sim Ge(p)$ p	$X \sim hy(x; N, M, n)$ N, M and n
Special case	When $n \to \infty$, $p \to 0$, use poission distribution	–	–	Put $r = 1$ in Negative binomial distribution	With $p = \frac{M}{N}$, $N \to \infty$ by $(x; N, M, n)$ approaches $b(n, p)$
MGF	$(q + pe^t)^n$	$e^{-\lambda(1-e^t)}$	$p^r(1 - qe^t)^{-r}$	$\frac{p}{1 - qe^t}$	–
Mean	np	λ	$\frac{rq}{p}$	$\frac{q}{p}$	$\frac{nM}{N}$
Variance	npq	λ	$\frac{rq}{p^2}$	$\frac{q}{p^2}$	$\frac{nM}{N}\left[\frac{(N-M)(N-n)}{N(N-1)}\right]$

Table 4.3

	Uniform Distribution	Normal Distribution	Gamma Distribution	Exponential Distribution	Chi square Distribution	Wel bull Distribution	Beta Distribution
PDF and Range	$\frac{1}{b-a}$ $a < x < b$	$\frac{1}{\sigma\sqrt{2}\pi}$ $e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty < x < \infty$	$\frac{1}{\sqrt{\alpha}\beta^{\alpha}}$ $x^{\alpha-1}e^{\frac{-x}{\beta}}$ $x > 0$	$\lambda e^{-\lambda x}$ $x > 0$	$\frac{1}{2^{\frac{r}{2}}\sqrt{\frac{r}{2}}}$ $x^{\frac{r}{2}-1}e^{-\frac{x}{2}}$ $x > 0$	$\alpha\beta x^{\beta-1}$ $e^{-\alpha x^{\beta}}$ $x > 0$	$\frac{1}{\beta(\alpha, \beta)}$ $x^{\alpha-1}(1-x)^{\beta-1}$ $0 < x < 1$
Symbol and Parameter(s)	$X \sim u(a, b)$ a and b	$X \sim n(\mu, \sigma^2)$ μ and σ^2	$X \sim G$ $(X; \alpha, \beta)$ α and β	$X \sim$ Expo (λ) λ	$X \sim X^2(r)$ r	$X \sim$ we $(x; \alpha, \beta)$ α and β	$X \sim$ Be $(x; \alpha, \beta)$ α and β
Special case		$\mu = 1, \sigma^2 = 0$ X is said to be standard normal variate		$\alpha = 1, \beta = \frac{1}{\lambda}$ Exponential distribution reduces to Gamma distribution	$\alpha = \frac{r}{2}\beta = 2$ Exponential distribution reduces to chi square distribution		
MGF	$\frac{e^{bt}-e^{at}}{t(b-a)}$	$e^{\mu t+\frac{\sigma^2 t^2}{2}}$	$\frac{1}{(1-\beta t)^{\alpha}}$	$\frac{\lambda}{\lambda-t}$	$\frac{1}{(1-2t)^{\frac{r}{2}}}$	–	–
Mean	$\frac{a+b}{2}$	μ	$\alpha\beta$	$\frac{1}{\lambda}$	r	$\frac{1}{\alpha^{\frac{1}{\beta}}}\sqrt{\frac{1}{\beta}+1}$	$\frac{\alpha}{\alpha+\beta}$
Variance	$\frac{(b-a)^2}{12}$	σ^2	$\alpha\beta^2$	$\frac{1}{\lambda^2}$	$2r$	$\frac{1}{\alpha^{\frac{2}{\beta}}}\left[\sqrt{\frac{2}{\beta}+1}-\left(\sqrt{\frac{1}{\beta}+1}\right)^2\right]$	$\frac{\alpha\beta}{(\alpha+\beta+1)^2(\alpha+\beta)}$

Exercise 4

1. In a certain city, it is claimed that 1 in 10 accidents is due to driver negligence. What is the probability that atleast 2 of 6 accidents are due to driver negligence?
2. A manufacture claims that only 2 percent of the items are defective. Assuming that this claim is true find the probabilities that among 20 items,
   (a) exactly 10 items are defective,
   (b) atleast 3 items are defective and
   (c) atmost 6 items are defective.
3. Find the probability of getting 9 heads in 10 tosses and 18 heads in 20 tosses of a fair coin.
4. Find the MGF and hence the mean and variance of a Bernoulli random variable and observe that letting $n = 1$, MGF of the binomial distribution becomes that of Bernoulli distribution.
5. Find the probabilities that
   (a) a family's fourth child is their first son,
   (b) a family's fifth child is their second or third daughter.
6. A die is thrown until 5 appears. What is the probability that it must be thrown atleast 6 times?
7. If probability is 0.75 that a target is destroyed in one shot, what is the probability that it would be destroyed on 6^{th} attempt?
8. If a fair coin is successively tossed, find the probability that a head first appears on the seventh trial. Also, find the probability that head appears third time on the seventh trial.
9. Suppose that two teams are playing a series of games, each of which is independently won by team A with probability 0.4. The winner of the series is the first team to win 4 games. Find the probability that atmost 7 games are played. Also, find the expected number of games are played.
10. If the life of a transistor tube X (in hours) is exponential with $\lambda = 1/180$, find the probability that the tube will last:
    (a) less than 36 hours,
    (b) between 36 and 90 hours and
    (c) more than 45 hours.
11. The daily consumption of milk in a city, in excess of 20,000 litres, is approximately distributed as a gamma variable with parameters $\alpha = 2, \beta = 10,000$. The city has a daily stock of 30,000 litres. What is the probability that the stock is insufficient on a particular day?
12. If X and Y are independent binomial variables with parameters 3, 0.7 and 4, 0.7 respectively. Find the probability that the sum of two random variables exceeds 3. Also, find the conditional probability that X takes on 3 given that the sum of X and Y is 4.
13. If $X \sim b(2, p)$, $Y \sim b(4, p)$ and $p(X \geq 1) = \frac{5}{9}$ find $p(Y \geq 1)$.

14. If $X \sim b(n, p)$, find $p(X = 4)$ and $p(X > 3)$ in each of the following cases.
    (a) $E(X) = 5$ and $E(X^2) = 29$,
    (b) $\mu = 6$ and $\sigma = 2$.
15. If X and Y are independent poisson variables such that $p(X = 1) = p(X = 2)$ and $p(Y = 2) = p(Y = 3)$, find the variance of $2X - 3Y$ and $X + 2Y$.
16. Prove that the linear combination of n independent normal variables is a normal variable. [Hint: Use MGF technique]
17. If $X \sim n(1, 9)$, $Y \sim n(2, 16)$, find the pdf of $U = X - Y$ and hence compute
    (a) $p[|U| < 3]$ and
    (b) $p[|U + 1| > 5]$.
18. If $X \sim n(1, 4)$, $Y \sim n(2, 3)$ and X and Y are independent, find the distribution of $U = 2X + 3Y$ and hence compute $p(U > 20)$.
19. If X follows exponential distribution with parameter λ and $p(X \leq 1) = p(X > 1)$. Find the mean and variance of X and also compute $p(X$ is atleast 2).
20. Identify the distribution whose MGF is
    (a) $M_X(t) = (0.7 + 0.3t)^5$, (b) $M_X(t) = \dfrac{1}{81}(2 + e^t)^4$,
    (c) $M_X(t) = e^{-3(1-e^t)}$, (d) $M_X(t) = e^{2t(1+t)}$,
    (e) $M_X(t) = \dfrac{1}{8\left(\frac{1}{2} - t\right)^3}$ and (f) $M_X(t) = \dfrac{1}{1 - \frac{t}{2}}$.
21. Find the MGF of $Y = 3X + 2$ for each of the case in problem (20).
22. A fair die is tossed 8 times. Find the probability of obtaining 5 and 6 exactly twice and the other numbers exactly once.
23. A car rental agency has 50 percent medium size cars, 30 percent luxury cars and 20 percent compact cars. In a sample of 10 cars, find the probability that 4 are medium size, 3 are luxury and 3 are compact cars.
24. The bivariate normal random variable (X, Y) have parameters $\mu_1 = 60, \mu_2 = 75, \sigma_1 = 6, \sigma_2 = 12$ and $\rho_{XY} = 0.55$. Find the following probabilities.
    (a) $p(65 \leq X \leq 75)$ and
    (b) $p(71 \leq Y \leq 80/X = 55)$.
25. Let X be the height of the husband and Y be the height of the wife in a certain population of married couples. Let X and Y have a bivariate normal distribution with parameter $\mu_1 = 174$ cms, $\mu_2 = 159$ cms $\sigma_1 = \sigma_2 = 6$ cm and $\rho = 0.6$. If the height of the husband is given as 189 cms, find the probability that his wife has a height between 158.4 and 177.6 cms.

26. The number of traffic accidents of a city in each month is assumed to be a poisson variable with $\lambda = 3$. What is the probability that there will be
    (a) exactly 5 traffic accidents,
    (b) atleast one accident in a certain month?
27. From a pack of 52 cards, 4 cards are drawn at random without replacement. What is the probability that two of them are aces?
28. If the probability of hitting a target is 0.4, what is the probability that the 10^{th} trial results in second hit?
29. A certain item is shipped in lots of 20. Five of them from a lot are taken at random and inspected. Each lot is accepted if none of them is defective. What is the probability that the lot is accepted when
    (a) there are 2 defects,
    (b) there are 10 defects in the lot?
30. If the average number of claims handled daily by an insurance company is 4, what proportion of days will have less than 2 claims? What is the probability that there will be 5 claims in exactly 2 of the next 5 days?
31. If the exponent of e of a bivariate normal density is

$$-\frac{1}{54}\left(x^2 + 4y^2 + 2xy + 2x + 8y + 4\right)$$

find σ_1, σ_2 and ρ given that $\mu_1 = 0$, $\mu_2 = -1$.
32. Find the distribution function of
    (a) uniform distribution,
    (b) exponential distribution and
    (c) weibull distribution.