
UNIT 7 CONTINUOUS PROBABILITY DISTRIBUTIONS AND EXACT SAMPLING DISTRIBUTIONS

Structure

- 7.0 Introduction
- 7.1 Objectives
- 7.2 Normal Distribution
- 7.3 Continuous Uniform Distribution
- 7.4 Introduction to χ^2 -Distribution
- 7.5 Introduction to t-Distribution
- 7.6 Introduction to F-Distribution
- 7.7 Summary
- 7.8 Solutions/Answers
- 7.9 References

7.0 INTRODUCTION

In the previous unit, we have studied standard discrete distributions. In this unit, we are going to discuss standard continuous univariate distributions. This unit introduces normal distribution, uniform distribution, chi-square distribution, t-distribution and F-distribution. Normal distribution has widespread applications. It is being used in almost all data-based research in the field of agriculture, trade, business, industry and society. For instance, normal distribution is a good approximation to the distribution of heights of randomly selected large number of students studying at the same level in a university. The continuous uniform distribution is a simple distribution that is useful in random phenomenon.

Sometimes statistics such as sample mean, sample proportion, sample variance, etc. may follow a particular sampling distribution. Based on this fact Prof. R.A. Fisher, Prof. G. Snedecor and some other statisticians worked in this area and found some exact sampling distributions which are often followed by some of the statistics. In this unit, we shall discuss some important sampling distributions such as χ^2 (read as chi-square), t and F. Generally, these sampling distributions are named on the name of the originator, that is, Fisher's F-distribution is named on Prof. R.A. Fisher.

7.1 OBJECTIVES

After studying this unit, you would be able to:

- introduce and explain the normal distribution;
- state various characteristics of the normal distribution;
- define continuous uniform and state its properties;
- describe the probability curve and properties of χ^2 -distribution;
- describe the probability curve and properties of t-distribution;
- describe the probability curve and properties of F-distribution;

- solve various practical problems based on the properties of various distributions.

7.2 NORMAL DISTRIBUTION

The concept of normal distribution was initially discovered by English mathematician Abraham De Moivre (1667-1754) in 1733. De Moivre obtained this continuous distribution as a limiting case of binomial distribution. His work was further refined by Pierre S. Laplace (1749-1827) in 1774. But the contribution of Laplace remained unnoticed for long till it was given concrete shape by Karl Gauss (1777-1855) who first made reference to it in 1809 as the distribution of errors in Astronomy. That is why the normal distribution is sometimes called Gaussian distribution. Though, normal distribution can be used as approximation to most of the other distributions, here we are going to discuss (without proof) its approximation to (i) binomial distribution and (ii) Poisson distribution.

Normal Distribution as a Limiting Case of Binomial Distribution

Normal distribution is a limiting case of binomial distribution under the following conditions:

- n , the number of trials, is indefinitely large i.e. $n \rightarrow \infty$;
- neither p (the probability of success) nor q (the probability of failure) is too close to zero.

Under these conditions, the binomial distribution can be closely associated by a normal distribution with standardised variable given by $Z = \frac{X - np}{\sqrt{npq}}$. The approximation becomes better with increasing n . In practice, the approximation is very good if both np and nq are greater than 5.

The normal distribution is symmetrical, and the curve of the distribution is mesokurtic, which is the main feature of the normal distribution.

Normal Distribution as a Limiting Case of Poisson Distribution

As we have discussed above that there is a relation between the binomial and normal distributions. It can, in fact, be shown that the Poisson distribution approaches a normal distribution with standardized variable given by $Z = \frac{X - \lambda}{\sqrt{\lambda}}$ as λ increases indefinitely.

Normal Distribution

Normal distribution is defined as follows:

Definition: A continuous random variable X is said to follow normal distribution with parameters μ ($-\infty < \mu < \infty$) and $\sigma^2 (>0)$ if it takes on any real value and its probability density function is given by

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

which may also be written as

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2 \right\}, -\infty < x < \infty.$$

Remark

- (i) The probability function represented by $f(x)$ may also be written as $f(x; \mu, \sigma^2)$.
- (ii) If a random variable X follows normal distribution with mean μ and variance σ^2 , then we may write, “ X is distributed to $N(\mu, \sigma^2)$ ” and is expressed as $X \sim N(\mu, \sigma^2)$.
- (iii) No continuous probability function, including the normal distribution, can be used to obtain the probability of occurrence of a particular value of the random variable. This is because such probability is very small, so instead of specifying the probability of taking a particular value by the random variable, we specify the probability of its lying within an interval.
- (iv) If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ is standard normal variate having mean ‘0’ and variance ‘1’. The values of mean and variance of standard normal variate are obtained as under, for which properties of expectation and variance are used.

$$\begin{aligned} \text{Mean of } Z \text{ i.e. } E(Z) &= E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma} [E(X-\mu)] \\ &= \frac{1}{\sigma} [E(X) - \mu] \\ &= \frac{1}{\sigma} [\mu - \mu] = 0 \quad [\because E(X) = \text{Mean of } X = \mu] \end{aligned}$$

$$\begin{aligned} \text{Variance of } Z \text{ i.e. } V(Z) &= V\left(\frac{X-\mu}{\sigma}\right) \\ &= \frac{1}{\sigma^2} [V(X-\mu)] = \frac{1}{\sigma^2} [V(X)] \\ &= \frac{1}{\sigma^2} (\sigma^2) \quad [\because \text{variance of } X \text{ is } \sigma^2] \\ &= 1. \end{aligned}$$

- (v) The probability density function of standard normal variate $Z = \frac{X-\mu}{\sigma}$ is given by $\phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$, $-\infty < z < \infty$.

This result can be obtained on replacing $f(x)$ by $\phi(z)$, x by z , μ by 0 and σ by 1 in the probability density function of normal variate X i.e. in

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

- (vi) The graph of the normal probability function $f(x)$ with respect to x is famous 'bell-shaped' curve. The top of the bell is directly above the mean μ . For large value of σ , the curve tends to flatten out and for small values of σ , it has a sharp peak as shown in (Fig. 7.1):

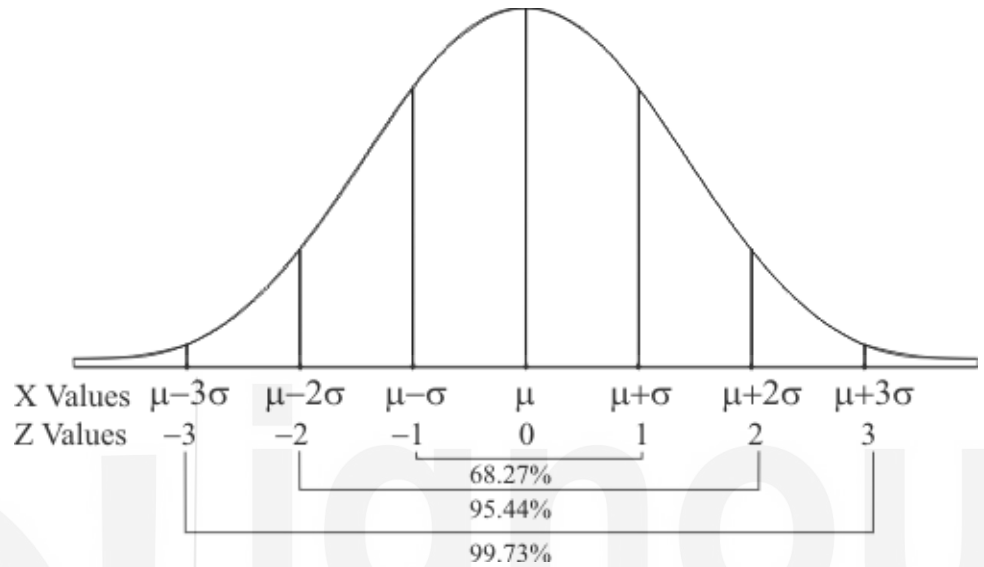


Fig. 7.1: The Normal Distribution

Normal distribution has various properties and large number of applications. It can be used as an approximation to most of the other distributions and hence is most important probability distribution in statistical analysis. Theory of estimation of population parameters and testing of hypotheses on the basis of sample statistics (to be discussed in the next unit) have also been developed using the concept of normal distribution as most of the sampling distributions tend to normality for large samples. Normal distribution has become widely and uncritically accepted on the basis of much practical work. As a result, it holds a central position in Statistics.

Let us now take some examples of writing the probability function of normal distribution when mean and variance are specified, and vice-versa.

Example 1: (i) If $X \sim N(40, 25)$ then write down the p.d.f. of X

(ii) If $X \sim N(-36, 20)$ then write down the p.d.f. of X

(iii) If $X \sim N(0, 2)$ then write down the p.d.f. of X

Solution: (i) Here we are given $X \sim N(40, 25)$

$\therefore$ in usual notations, we have

$$\mu = 40, \sigma^2 = 25 \Rightarrow \sigma = \pm\sqrt{25}$$

$$\Rightarrow \sigma = 5 \quad [\because \sigma > 0 \text{ always}]$$

Now, the p.d.f. of random variable X is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$= \frac{1}{5\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-40}{5}\right)^2}, \quad -\infty < x < \infty$$

(ii) Here we are given $X \sim N(-36, 20)$.

$\therefore$ in usual notations, we have

$$\mu = -36, \sigma^2 = 20 \Rightarrow \sigma = \sqrt{20}$$

Now, the p.d.f. of random variable X is given by

$$\begin{aligned} f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \\ &= \frac{1}{\sqrt{20}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-(-36)}{\sqrt{20}}\right)^2} = \frac{1}{\sqrt{40\pi}} e^{-\frac{1}{40}(x+36)^2} \\ &= \frac{1}{2\sqrt{10\pi}} e^{-\frac{1}{40}(x+36)^2}, \quad -\infty < x < \infty \end{aligned}$$

(iii) Here we are given $X \sim N(0, 2)$.

$\therefore$ in usual notations, we have

$$\mu = 0, \sigma^2 = 2 \Rightarrow \sigma = \sqrt{2}$$

Now, the p.d.f. of random variable X is given by

$$\begin{aligned} f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{\sqrt{2}}\right)^2} \\ &= \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}x^2}, \quad -\infty < x < \infty \end{aligned}$$

Example 2: Below, in each case, there is a p.d.f. of a normally distributed random variable. Obtain the parameters (mean and variance) of the variable.

$$(i) f(x) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-46}{6}\right)^2}, \quad -\infty < x < \infty$$

$$(ii) f(x) = \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{32}(x-60)^2}, \quad -\infty < x < \infty$$

Solution: (i) $f(x) = \frac{1}{6\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-46}{6}\right)^2}, \quad -\infty < x < \infty$

Comparing it with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

we have

$$\mu = 46, \quad \sigma = 6$$

$$\therefore \text{Mean} = \mu = 46, \text{ variance} = \sigma^2 = 36$$

$$\begin{aligned} \text{(ii) } f(x) &= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{32}(x-60)^2}, \quad -\infty < x < \infty \\ &= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-60}{4}\right)^2} \end{aligned}$$

Comparing it with,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

we get

$$\mu = 60, \quad \sigma = 4$$

$$\therefore \text{Mean} = \mu = 60, \text{ variance} = \sigma^2 = 16$$

Now, we are going to state some important properties of Normal distribution in the next section.

Main Characteristics of Normal Distribution

The normal probability distribution with mean μ and variance σ^2 has the following properties:

- i) The curve of the normal distribution is bell-shaped as shown in Fig. 7.1.
- ii) The curve of the distribution is completely symmetrical about $x = \mu$, i.e. if we fold the curve at $x = \mu$, both the parts of the curve are the mirror images of each other.
- iii) For normal distribution, **Mean = Median = Mode**
- iv) $f(x)$, being the probability, can never be negative and hence no portion of the curve lies below x-axis.
- v) Though x-axis becomes closer and closer to the normal curve as the magnitude of the value of x goes towards ∞ or $-\infty$, yet it never touches it.
- vi) Normal curve has only one mode.
- vii) Central moments of Normal distribution are

$$\mu_1 = 0, \mu_2 = \sigma^2, \mu_3 = 0, \mu_4 = 3\sigma^4$$

The distribution is symmetrical and curve is always mesokurtic.

Note: Not only μ_1 and μ_3 are 0 but all the odd order central moments are zero for a normal distribution.

- viii) For normal curve,

$$Q_3 - \text{Median} = \text{Median} - Q_1$$

i.e. First and third quartiles of normal distribution are equidistant from median.

ix) Quartile Deviation (Q.D.) = $\frac{Q_3 - Q_1}{2}$ is approximately equal to $\frac{2}{3}$ of the standard deviation.

x) Mean deviation is approximately equal to $\frac{4}{5}$ of the standard deviation.

xi) If $X_1, X_2, \dots, X_n$ are independent normal variables with means $\mu_1, \mu_2, \dots, \mu_n$ and variances $\sigma_1^2, \sigma_2^2, \dots, \sigma_n^2$ respectively, then the linear combination $a_1X_1 + a_2X_2 + \dots + a_nX_n$ of $X_1, X_2, \dots, X_n$ is also a normal variable with mean

$$a_1\mu_1 + a_2\mu_2 + \dots + a_n\mu_n \text{ and variance } a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + \dots + a_n^2\sigma_n^2.$$

xii) Particularly, sum or difference of two independent normal variates is also a normal variate. If X and Y are two independent normal variates with means μ_1, μ_2 and variances σ_1^2, σ_2^2 , then

$$X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \text{ and } X - Y \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2).$$

Also, if $X_1, X_2, \dots, X_n$ are independent variates each distributed as:

$$N(\mu, \sigma^2), \text{ then their mean } \bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right).$$

xiii) Area property:

$$P[\mu - \sigma < X < \mu + \sigma] = \int_{\mu - \sigma}^{\mu + \sigma} f(x) dx = 0.6827,$$

$$\text{Or } P[-1 < Z < 1] = \int_{-1}^1 \phi(z) dz = 0.6827,$$

$$P[\mu - 2\sigma < X < \mu + 2\sigma] = \int_{\mu - 2\sigma}^{\mu + 2\sigma} f(x) dx = 0.9544,$$

$$\text{Or } P[-2 < Z < 2] = \int_{-2}^2 \phi(z) dz = 0.9544, \text{ and}$$

$$P[\mu - 3\sigma < X < \mu + 3\sigma] = \int_{\mu - 3\sigma}^{\mu + 3\sigma} f(x) dx = 0.9973.$$

$$\text{Or } P[-3 < Z < 3] = \int_{-3}^3 \phi(z) dz = 0.9973.$$

Let us now take up some problems based on properties of Normal Distribution in the next section.

Example 3: If X_1 and X_2 are two independent variates each distributed as $N(0, 1)$, then write the distribution of (i) $X_1 + X_2$. (ii) $X_1 - X_2$.

Solution: We know that, if X_1 and X_2 are two independent normal variates s.t.

$$X_1 \sim N(\mu_1, \sigma_1^2) \text{ and } X_2 \sim N(\mu_2, \sigma_2^2)$$

then

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2), \text{ and}$$

$$X_1 - X_2 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \quad [\text{See Property xii (Section 7.3)}]$$

Here, $X_1 \sim N(0,1)$, $X_2 \sim N(0,1)$

$$\therefore \text{ i) } X_1 + X_2 \sim N(0+0, 1+1)$$

$$\text{ i.e. } X_1 + X_2 \sim N(0, 2), \text{ and}$$

$$\text{ ii) } X_1 - X_2 \sim N(0-0, 1+1)$$

$$\text{ i.e. } X_1 - X_2 \sim N(0, 2)$$

Example 4: If $X \sim N(30, 25)$, find the mean deviation about mean.

Solution: Here $\mu = 30$, $\sigma^2 = 25 \Rightarrow \sigma = 5$.

$$\therefore \text{ Mean deviation about mean} = \sqrt{\frac{2}{\pi}} \sigma = \sqrt{\frac{2}{\pi}} \cdot 5 = 5\sqrt{\frac{2}{\pi}}$$

You can now try some exercises based on the properties of normal distribution which you have studied in the present unit.

Check Your Progress 1

1) Write down the p.d.f. of r. v. X in each of the following cases:

$$\text{ (i) } X \sim N\left(\frac{1}{2}, \frac{4}{9}\right)$$

$$\text{ (ii) } X \sim N(-40, 16)$$

.....

.....

2) Below, in each case, is given the p.d.f. of a normally distributed random variable. Obtain the parameters (mean and variance) of the variable.

$$\text{ (i) } f(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{x^2}{8}}, \quad -\infty < x < \infty$$

$$\text{ (ii) } f(x) = \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}(x-2)^2}, \quad -\infty < x < \infty$$

.....

.....

3) If X_1 and X_2 are two independent normal variates with means 30, 40 and variances 25, 35 respectively. Find the mean and variance of

$$\text{ i) } X_1 + X_2$$

$$\text{ ii) } X_1 - X_2$$

.....

.....

7.3 CONTINUOUS UNIFORM DISTRIBUTION

The uniform (or rectangular) distribution is a very simple distribution. It provides a useful model for a few random phenomena like having random number from the interval $[0, 1]$, then one is thinking of the value of a uniformly distributed random variable over the interval $[0, 1]$.

Definition: A random variable X is said to follow a continuous uniform (rectangular) distribution over an interval (a, b) if its probability density function is given by

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

The distribution is called uniform distribution since it assumes a constant (uniform) value for all x in (a, b) . If we draw the graph of $y = f(x)$ over x -axis and between the ordinates $x = a$ and $x = b$ (say), it describes a rectangle as shown in Fig. 7.2

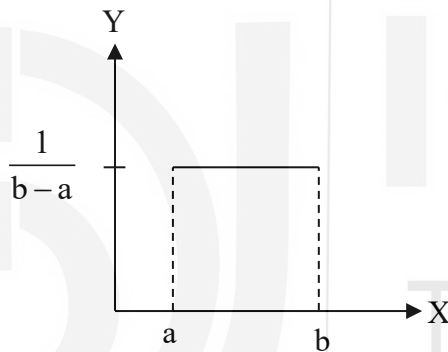


Fig. 7.2: Graph of a uniform function

A uniform variate X on the interval (a, b) is written as $X \sim U[a, b]$

Cumulative Distribution Function

The cumulative distribution function of the uniform random variate over the interval (a, b) is given by:

$$F(x) = \begin{cases} 0 & \text{for } x \leq a \\ \frac{x-a}{b-a} & \text{for } a < x < b \\ 1 & \text{for } x \geq b \end{cases}$$

On plotting its graph, we have

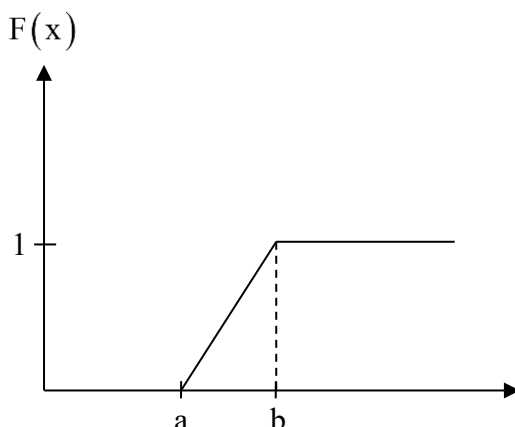


Fig. 7.3: Graph of distribution function

Mean and Variance of Uniform Distribution

$$\text{Mean} = \frac{a+b}{2} \text{ and Variance} = \frac{(b-a)^2}{12}.$$

Let us now take up some examples of continuous uniform distribution.

Example 5: If X is uniformly distributed with mean 2 and variance 12, find $P[X < 3]$.

Solution: Let $X \sim U[a, b]$

$\therefore$ probability density function of X is

$$f(x) = \frac{1}{b-a}, \quad a < x < b.$$

Now as Mean = 2

$$\Rightarrow \frac{a+b}{2} = 2$$

$$\Rightarrow a + b = 4 \quad \dots (1)$$

Variance = 12

$$\Rightarrow \frac{(b-a)^2}{12} = 12$$

$$\Rightarrow (b-a)^2 = 144$$

$$\Rightarrow b - a = \pm 12$$

$$\Rightarrow b - a = 12 \quad \dots (2)$$

$\left[\begin{array}{l} \because b - a = -12, \text{ being negative is} \\ \text{rejected as } b \text{ should be greater than } a \\ \Rightarrow b - a \text{ should be positive} \end{array} \right]$

Adding (1) and (2), we have

$$2b = 16$$

$$\Rightarrow b = 8 \text{ and hence } a = 4 - 8 = -4$$

$$\therefore f(x) = \frac{1}{b-a} = \frac{1}{8-(-4)} = \frac{1}{12} \text{ for } -4 < x < 8.$$

$$\begin{aligned} \text{Thus, the desired probability} &= P[X < 3] = \int_{-4}^3 \frac{1}{12} dx = \frac{1}{12} \int_{-4}^3 1 dx = \frac{1}{12} [x]_{-4}^3 \\ &= \frac{1}{12} [3 - (-4)] = \frac{7}{12}. \end{aligned}$$

Example 6: Metro trains are scheduled every 5 minutes at a certain station. A person comes to the station at a random time. Let the random variable X count the number of minutes he/she has to wait for the next train. Assume X has a uniform distribution over the interval $(0, 5)$. Find the probability that he/she has to wait at least 3 minutes for the train.

Solution: As X follows uniform distribution over the interval $(0, 5)$,

∴ probability density function of X is

$$f(x) = \frac{1}{b-a} = \frac{1}{5-0} = \frac{1}{5}, 0 < x < 5$$

Thus, the desired probability

$$\begin{aligned} P[X \geq 3] &= \int_3^5 f(x) dx = \int_3^5 \frac{1}{5} dx = \frac{1}{5} \int_3^5 (1) dx \\ &= \frac{1}{5} [x]_3^5 = \frac{1}{5} (5-3) = \frac{2}{5} = 0.4 \end{aligned}$$

Check Your Progress 2

1) Suppose that X is uniformly distributed over $(-a, a)$. Determine 'a' so that

i) $P[X > 4] = \frac{1}{3}$

ii) $P[X < 1] = \frac{3}{4}$

2) A random variable X has a uniform distribution over $(-2, 2)$. Find k for which $P[X > k] = \frac{1}{2}$.

7.4 INTRODUCTION TO χ^2 -DISTRIBUTION

Before describing the chi-square (χ^2) distribution, first, we will discuss a very useful concept “degrees of freedom”. The exact sampling distributions are described with the help of degrees of freedom.

Degrees of Freedom (df)

The term degree of freedom (df) is related to the independence of sample observations. In general, the number of degree of freedom is the total number of observations minus the number of independent constraints or restrictions imposed on the observations.

For a sample of n observations, if there are k restrictions among observations ($k < n$), then the degrees of freedom will be $(n-k)$.

For example, suppose there are two observations X_1, X_2 and a restriction (condition) is imposed that their total should be equal to 100. Then we can arbitrarily assign value to any one of these two as 30, 98, 52, etc. but the value

of rest is automatically determined by a simple adjustment like $X_2 = (100 - X_1)$. Therefore, we can say that only one observation is independent or there is one degree of freedom.

Consider another example, suppose there are 50 observations as $X_1, X_2, \dots, X_{50}$ and the restriction (condition) are imposed on these, like:

$$(i) \sum_{i=1}^{50} X_i = 2000 \quad (ii) \sum_{i=1}^{50} X_i^2 = 10000 \quad (iii) X_5 = 2X_4$$

The restriction (i) reveals that the sum of observations equal to 2000 (ii) reveals that the sum of squares of observations equal to 10000 and restriction (iii) implies that the fifth observation is double of the fourth observation.

With these three restrictions (i), (ii) and (iii), there shall be only 47 observations independently chosen and rest three could be obtained by solving equations or by adjustments. The degrees of freedom of this set of observations of size 50 is now $(50 - 3) = 47$.

Chi-square Distribution

The chi-square distribution is first discovered by Helmer in 1876 and later independently explained by Karl- Pearson in 1900. The chi-square distribution was discovered mainly as a measure of goodness of fit in case of frequency.

If a random sample $X_1, X_2, \dots, X_n$ of size n is drawn from a normal population having mean μ and variance σ^2 then the sample variance can be defined as

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

or

$$\sum_{i=1}^n (X_i - \bar{X})^2 = (n-1)S^2 = vS^2$$

where, $v = n - 1$ and the symbol v read as 'nu'.

Generally, v symbol is used to represent the general degrees of freedom. Its value may be $n, n - 1$, etc.

Thus, the variate $\chi^2 = \frac{vS^2}{\sigma^2}$, which is the ratio of sample variance multiplied by its degrees of freedom and the population variance follows the χ^2 -distribution with v degrees of freedom.

Properties of χ^2 -distribution

In the previous section, we have discussed the χ^2 -distribution and its probability density function. Now, in this section, we shall discuss some important properties of χ^2 -distribution as given below:

1. The probability curve of the chi-square distribution lies in the first quadrant because the range of χ^2 -variate is from 0 to ∞ .
2. Chi-square distribution has only one parameter n , that is, the degrees of freedom.
3. Chi-square probability curve is highly positively skewed.
4. Chi-square distribution is a uni-modal distribution, that is, it has a single mode.

5. The mean and variance of the chi-square distribution with n df are n and $2n$ respectively.

After deliberating the properties of χ^2 -distribution above, some of these are discussed in detail in the subsequent sub-sections.

Probability Curve of χ^2 -distribution

The probability distribution of all possible values of a statistic is known as the sampling distribution of that statistic. The shape of the probability distribution of a statistic can be shown by the probability curve. In this sub-section, we will discuss the probability curve of chi-square distribution.

A rough sketch of the probability curve of chi-square distribution is shown in Fig. 7.4 at different values of degrees of freedom as $n = 1, 4, 10$ and 22 .

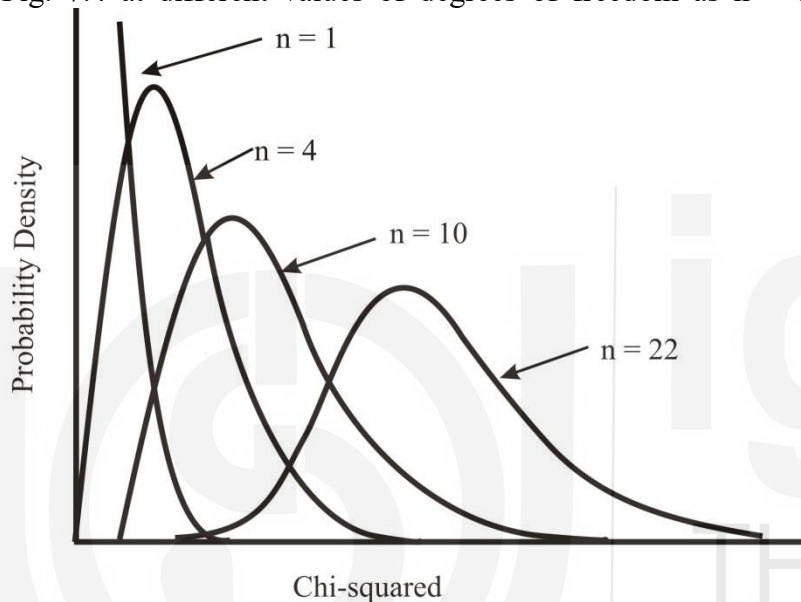


Fig. 7.4: Chi-square probability curves for $n = 1, 4, 10$ and 22

After looking at the probability curves of χ^2 -distribution for $n = 1, 4, 10$ and 22 , one can understand that probability curve takes shape of inverse J for $n = 1$. The probability curve of chi-square distribution gets skewed more and more to the right as n becomes smaller and smaller. It becomes more and more symmetrical as n increases because as n tends to ∞ , the χ^2 -distribution becomes a normal distribution.

Mean and Variance of χ^2 -distribution

The mean and variance of χ^2 -distribution are derived with the help of the moment about origin. The mean and variance of the chi-square distribution with n df are given by

$$\text{Mean} = n \text{ and Variance} = 2n$$

Hence, we have observed that the variance is twice the mean of the chi-square distribution.

Example 7: What are the mean and variance of the chi-square distribution with 5 degrees of freedom?

Solution: We know that the mean and variance of the chi-square distribution with n degrees of freedom are

$$\text{Mean} = n \text{ and Variance} = 2n$$

In our case, $n = 5$, therefore, Mean = 5 and Variance = 10.

Applications of χ^2 -Distribution

The applications of chi-square distribution are very wide in Statistics. Some of them are listed below. The chi-square distribution is used:

1. To test the hypothetical value of population variance.
2. To test the goodness of fit, that is, to judge whether there is a discrepancy between theoretical and experimental observations.
3. To test the independence of two attributes, that is, to judge whether the two attributes are independent.

Try the following check your progress to make sure that you have understood the chi-square distribution.

Check Your Progress 3

- 1) What are the mean and variance of the chi-square distribution with 10 degrees of freedom?

.....
.....

- 2) List the applications of the chi-square distribution.

.....
.....

7.5 INTRODUCTION TO t-DISTRIBUTION

The t-distribution was discovered by W.S. Gosset in 1908. He was better known by the pseudonym ‘Student’ and hence t-distribution is called ‘Student’s t-distribution’.

If a random sample $X_1, X_2, \dots, X_n$ of size n is drawn from a normal population having mean μ and variance σ^2 then we know that the sample mean $\bar{X}$ is distributed normally with mean μ and variance σ^2 / n , that is, if $X_i \sim N(\mu, \sigma^2)$

then $\bar{X} \sim N(\mu, \sigma^2 / n)$, and also the variate

$$Z = \frac{\bar{X} - \mu}{\sigma / \sqrt{n}}$$

is distributed normally with mean 0 and variance 1, i.e. $Z \sim N(0, 1)$.

In general, the standard deviation σ is not known and in such a situation the only alternative left is to estimate the unknown σ^2 . The value of sample variance (S^2) is used to estimate the unknown σ^2 where,

$$S^2 = \frac{1}{(n-1)} \sum_{i=1}^n (X_i - \bar{X})^2$$

Thus, in this case, the variate $\frac{\bar{X} - \mu}{S / \sqrt{n}}$ is not normally distributed, whereas it

follows t-distribution with $(n-1)$ df i.e.

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{(n-1)} \quad \dots (3)$$

$$\text{where } S^2 = \frac{1}{(n-1)} \sum_{i=1}^n (X_i - \bar{X})^2$$

The t-variate is a widely used variable and its distribution is called student's t-distribution on the pseudonym name 'Student' of W.S. Gosset.

After describing the t-distribution, we try to calculate the value of t-variate as in the given example.

Example 8: The life of light bulbs manufactured by company A is known to be normally distributed. The CEO of the company claims that an average lifetime of the light bulbs is 300 days. A researcher randomly selects 25 bulbs for testing the lifetime and he observed the average lifetime of the sampled bulbs is 290 days with standard deviation of 50 days. Calculate the value of t-variate.

Solution: Here, we are given that

$$\mu = 300, n = 25, \bar{X} = 290 \text{ and } S = 50$$

The value of t-variate can be calculated by the formula given below

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Therefore, we have

$$t = \frac{290 - 300}{50/\sqrt{25}} = \frac{-10}{10} = -1$$

Properties of t-Distribution

In the previous section, we have discussed the t-distribution briefly. Now, we shall discuss some of the important properties of the t-distribution. The t-distribution has the following properties:

1. The t-distribution is a uni-modal distribution, that is, t-distribution has a single mode.
2. The mean and variance of the t-distribution with n df are zero and $\frac{n}{n-2}$ respectively for $n > 2$.
3. The probability curve of t-distribution is similar in shape to the standard normal distribution and is symmetric about $t = 0$ line but flatter than the normal curve.
4. The probability curve is bell-shaped and asymptotic to the horizontal axis.

The properties deliberated above are some of the important properties of the t-distribution. We shall discuss some of these properties such as probability curve, mean and variance of the t-distribution in detail in the subsequent sub-sections.

Probability Curve of t-distribution

The probability curve of t-distribution is bell-shaped and symmetric about $t = 0$ line. The probability curves of t-distribution are shown in Fig. 7.5 at two different values of degrees of freedom, $n = 4$ and $n = 12$.

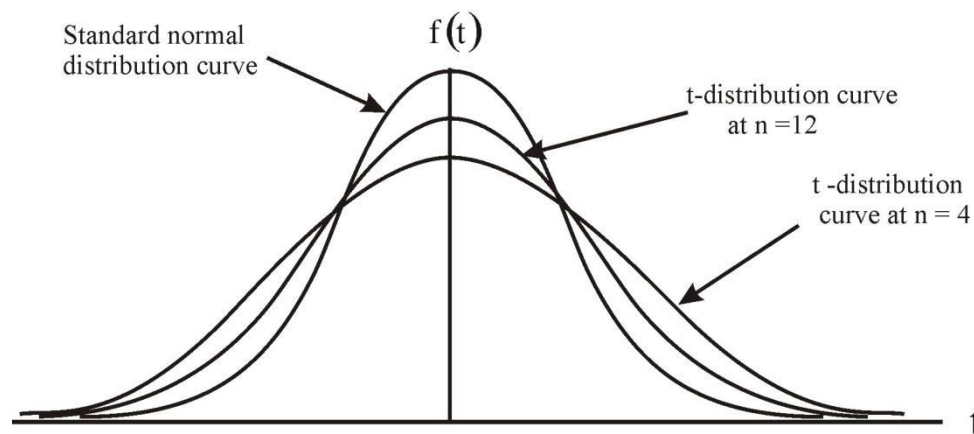


Fig. 7.5: Probability curves for t-distribution at $n = 4, 12$ along with standard normal curve

In the figure given above, we have drawn the probability curves of t-distribution at two different values of degrees of freedom with the probability curve of standard normal distribution. By looking at the figure, one can easily understand that the probability curve of t-distribution is similar in shape to that of normal distribution and asymptotic to the horizontal-axis whereas it is flatter than the standard normal curve. The probability curve of the t-distribution tends to the normal curve as the value of n increases. Therefore, for a sufficiently large value of sample size $n(> 30)$, the t-distribution tends to the normal distribution.

Mean and Variance of t-distribution

The probability curve shows that t-distribution is symmetrical about $t = 0$ line. Thus, mean of the t-distribution is zero.

The mean and variance of the t-distribution with n df are given by

$$\text{Mean} = 0 \text{ and Variance} = \frac{n}{n-2} \text{ for } n > 2$$

Applications of t-Distribution

After learning the important properties of t-distribution and its mean and variance. Now, you may be interested to know the applications of the t-distribution. The t-distribution has a wide number of applications in Statistics. The t-distribution is used:

1. To test the hypothesis about the population mean.
2. To test the hypothesis about the difference of two population means of two normal populations.
3. To test the hypothesis that the population correlation coefficient is zero.

Check Your Progress 4

- 1) If the scores on an IQ test of the students of a class are assumed to be normally distributed with mean of 60. From the class, 15 students are randomly selected and an IQ test of a similar level is conducted. The average test score and standard deviation of test scores in the sample group are found to be 65 and 12 respectively. Calculate the value of t-variate.

.....

.....

.....

.....

- 2) What are the mean and variance of t-distribution with 20 degrees of freedom?

.....

- 3) Write any three applications of t-distribution.

.....

7.6 INTRODUCTION TO F-DISTRIBUTION

F-distribution was introduced by Prof. R. A. Fisher and defined as the ratio of two independent chi-square variates when divided by their respective degrees of freedom. If we draw a random sample $X_1, X_2, \dots, X_{n_1}$ of size n_1 from a normal population with mean μ_1 and variance σ_1^2 and another independent random sample $Y_1, Y_2, \dots, Y_{n_2}$ of size n_2 from another normal population with mean μ_2 and variance σ_2^2 respectively. Also, $v_1 S_1^2 / \sigma_1^2$ is distributed as chi-square variate with v_1 df i.e.

$$\chi_1^2 = \frac{v_1 S_1^2}{\sigma_1^2} \sim \chi_{(v_1)}^2 \quad \dots (4)$$

$$\text{where, } v_1 = n_1 - 1, \bar{X} = \frac{1}{n_1} \sum_{i=1}^{n_1} X_i \text{ and } S_1^2 = \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_i - \bar{X})^2$$

Similarly, $v_2 S_2^2 / \sigma_2^2$ is distributed as chi-square variate with v_2 df i.e.

$$\chi_2^2 = \frac{v_2 S_2^2}{\sigma_2^2} \sim \chi_{(v_2)}^2 \quad \dots (5)$$

$$\text{where, } v_2 = n_2 - 1, \bar{Y} = \frac{1}{n_2} \sum_{i=1}^{n_2} Y_i \text{ and } S_2^2 = \frac{1}{n_2 - 1} \sum_{i=1}^{n_2} (Y_i - \bar{Y})^2$$

Now, if we take the ratio of the above chi-square variates given in equations (4) and (5), then we get

$$\begin{aligned} \frac{\chi_1^2}{\chi_2^2} &= \frac{v_1 S_1^2 / \sigma_1^2}{v_2 S_2^2 / \sigma_2^2} \\ \Rightarrow \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2} &= \frac{\chi_1^2 / v_1}{\chi_2^2 / v_2} \sim F_{(v_1, v_2)} \quad \dots (6) \end{aligned}$$

By observing the above form given in equation (6) we reveal that the ratio of two independent chi-square variates when divided by their respective degrees of freedom follows F-distribution where, v_1 and v_2 are called the degrees of freedom of F-distribution.

Now, if variances of both the populations are equal i.e. $\sigma_1^2 = \sigma_2^2$ then F-variate is written in the form of ratio of two sample variances i.e.

$$F = \frac{S_1^2}{S_2^2} \sim F_{(v_1, v_2)} \quad \dots (4)$$

After describing the F-distribution, we try to calculate the value of F-variate as in the given example.

Example 9: A statistician selects 7 women randomly from the population of women, and 12 men from a population of men. The table given below shows the standard deviation of each sample and population:

Population	Population Standard Deviation	Sample Standard Deviation
Women	40	45
Men	80	75

Compute the value of F-variate.

Solution: Here, we are given that

$$n_1 = 7, n_2 = 12, \sigma_1 = 40, \sigma_2 = 80, S_1 = 45, S_2 = 75$$

The value of F-variate can be calculated by the formula given below

$$F = \frac{S_1^2 / \sigma_1^2}{S_2^2 / \sigma_2^2}$$

where, S_1^2 & S_2^2 are the sample variances.

Therefore, we have

$$F = \frac{(45)^2 / (40)^2}{(75)^2 / (80)^2} = \frac{1.27}{0.88} = 1.93$$

For the above calculation, the degrees of freedom v_1 for women's data are $7 - 1 = 6$ and the degrees of freedom v_2 for men's data are $12 - 1 = 11$.

After introducing the F-distribution, one may be interested in knowing the properties of this distribution. Now, we shall discuss some of the important properties of F-distribution.

Properties of F-Distribution

The F-distribution has wide properties in Statistics. Some of them are as follow:

1. The probability curve of F-distribution is positively skewed. The curve becomes highly positive skewed when v_2 is smaller than v_1 .
2. F-distribution curve extends on abscissa from 0 to ∞ .
3. F-distribution is a uni-modal distribution, that is, it has a single mode.
4. The square of t-variate with v df follows F-distribution with 1 and v degrees of freedom.

5. The mean of F-distribution with (v_1, v_2) df is $\frac{v_2}{v_2 - 2}$ for $v_2 > 2$.
6. The variance of F-distribution with (v_1, v_2) df is

$$\frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)} \text{ for } v_2 > 4.$$

7. If we interchange the degrees of freedom v_1 and v_2 then there exists a very useful relation as

$$F_{(v_1, v_2), (1-\alpha)} = \frac{1}{F_{(v_2, v_1), \alpha}}$$

Probability Curve of F-distribution

F-variate is the ratio of two sample variances when population variances are equal and follows F-distribution with v_1 and v_2 df i.e.

$$F = \frac{S_1^2}{S_2^2} \sim F_{(v_1, v_2)} \quad \dots (6)$$

where, S_1^2 and S_2^2 are the sample variances of two normal populations. As you have studied that the F-variate varies from 0 to ∞ , therefore, it is always positive so probability curve of F-distribution wholly lies in the first quadrant. These v_1 and v_2 are the degrees of freedom and are called the parameters of F-distribution. Hence, the shape of the probability curve F-distribution depends on v_1 and v_2 . The probability curves of F-distribution for various pairs of degrees of freedom (v_1, v_2) i.e. (5, 5), (5, 20) and (20, 5) are shown in Fig. 7.6 given below:

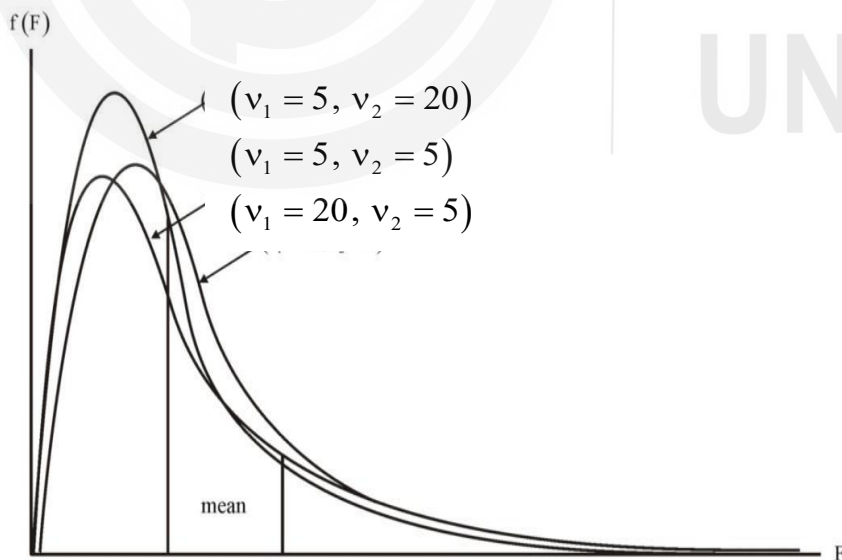


Fig. 7.6: Probability curves of F-distribution for (5, 5), (5, 20) and (20, 5) degrees of freedom.

After looking at the probability curves of F-distribution drawn above, we can observe that the probability curve of the F-distribution is a uni-model curve. And by increasing in the first degrees of freedom from $v_1 = 5$ to $v_1 = 20$ the mean of the distribution (shown by the vertical line) does not change but probability

curve shifted from the tail to the centre of the distribution whereas increasing in the second degrees of freedom from $v_2 = 5$ to $v_2 = 20$ the mean of the distribution (shown by the vertical line) decrease and the probability curve shifted from the tail to the centre of the distribution. One can also get an idea about the skewness of the F-distribution, we have observed from the probability curve of the F-distribution that it is positively skewed curve and it becomes very highly positive skewed if v_2 becomes small.

Mean and Variance of F-distribution

The mean and variance of F-distribution are given below:

Mean

$$= \frac{v_2}{v_2 - 2} \quad \text{for } v_2 > 2$$

Variance

$$= \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)} \quad \text{for } v_2 > 4$$

From the above derivation of mean and variance, we can conclude that mean of F-distribution is independent of v_1 .

Applications of F-Distribution

Though F-distribution has many applications some of them are listed here.

The F-distribution has the following applications:

- F-distribution is used to test the hypothesis about equality of the variances of two normal populations.
- F-distribution is used to test the hypothesis about multiple correlation coefficients.
- F-distribution is used to test the hypothesis about correlation ratio.
- F-distribution is used to test the equality of means of k-populations, when one characteristic of the population is considered i.e. F-distribution is used in one-way analysis of variance.
- F-distribution is used to test the equality of k-population means for two characteristics at a time i.e. F-distribution is used in two-way analysis of variance.

Relation Between t , χ^2 and F-Distributions

After describing the exact sampling distributions χ^2 , t & F , we shall explain the relationship between t & F -distributions and χ^2 & F -distributions. These relations can be described in terms of the following two statements:

Statement 1: If a variate t follows Student's t -distribution with v df, then square of t follows F-distribution with $(1, v)$ df i.e. if $t \sim t_{(v)}$ then $t^2 \sim F(1, v)$.

Statement 2: If F-variate follows F-distribution with (v_1, v_2) df and if

$v_2 \rightarrow \infty$ then $\chi^2 = v_1 F$ follows a χ^2 -distribution with v_1 df.

The proofs of these statements are beyond the scope of this course.

Check Your Progress 5

- 1) For a survey, 15 students are selected randomly from class A and 10 students are selected randomly from class B. At the stage of the analysis of the sample data, the following information is available:

Class	Population Standard Deviation	Sample Standard Deviation
Class A	65	60
Class B	45	50

Calculate the value of F-variate.

.....

- 2) Write any five properties of F-distribution.

.....

- 3) What are the mean and variance of F-distribution with $v_1 = 5$ and $v_2 = 10$ degrees of freedom?

.....

- 4) Write four applications of F-distribution.

.....

Tabulated Values Of Z, t, χ^2 and F

The area or probability of normal, t, χ^2 and F-distributions can be calculated with the help of tables (See Appendix). You can also find the area or probability of a variate using tables. You can refer to the references for the same.

We now end this unit by giving a summary of what we have covered in it.

7.7 SUMMARY

The following main points have been covered in this unit:

- 1) A continuous random variable X is said to follow normal distribution with parameters μ ($-\infty < \mu < \infty$) and $\sigma^2(>0)$ if it takes on any real value and its probability density function is given by

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

- 2) If $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma}$ is standard normal variate.
- 3) The curve of the normal distribution is bell-shaped and is completely symmetrical about $x = \mu$.
- 4) For normal distribution, Mean = Median = Mode.
- 5) Sum of independent normal variables is also a normal variable.
- 6) A random variable X is said to follow a continuous uniform (rectangular) distribution over an interval (a, b) if its probability density function is given by

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

- 7) For **continuous uniform distribution**, **Mean is $(a+b)/2$** and **variance is $(b-a)^2/12$**
- 8) The unit introduced χ^2 -distribution, its probability curve, its properties and applications.
- 9) The unit introduced t-distribution, its probability curve, its properties and applications.
- 10) The unit introduced F-distribution, its probability curve, its properties and applications.

7.8 SOLUTIONS/ANSWERS

Check Your Progress 1

- 1) (i) Here we are given $X \sim N\left(\frac{1}{2}, \frac{4}{9}\right)$

$\therefore$ in usual notations, we have

$$\mu = \frac{1}{2}, \quad \sigma^2 = \frac{4}{9} \quad \Rightarrow \quad \sigma = \frac{2}{3}$$

Now, p.d.f. of r.v. X is given by

$$\begin{aligned} f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty \\ &= \frac{1}{\frac{2}{3}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-1/2}{2/3}\right)^2} \\ &= \frac{3}{2\sqrt{2\pi}} e^{-\frac{9}{2}\left(\frac{2x-1}{4}\right)^2}, \quad -\infty < x < \infty \end{aligned}$$

- (ii) Here we are given $X \sim N(-40, 16)$

∴ in usual notations, we have

$$\mu = -40, \quad \sigma^2 = 16 \Rightarrow \sigma = 4$$

Now, p.d.f. of r.v. X is given by

$$\begin{aligned} f(x) &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}, \quad -\infty < x < \infty \\ &= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-(-40)}{4}\right)^2} \\ &= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x+40}{4}\right)^2}, \quad -\infty < x < \infty \end{aligned}$$

$$2) (i) f(x) = \frac{1}{2\sqrt{2\pi}} e^{-\frac{x^2}{8}}, \quad -\infty < x < \infty$$

$$\begin{aligned} &= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{4}} \\ &= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{2}\right)^2} \quad \dots(1), -\infty < x < \infty \end{aligned}$$

Comparing (1) with,

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

we get

$$\mu = 0, \quad \sigma = 2$$

$$\therefore \text{Mean} = \mu = 0 \text{ and variance} = \sigma^2 = (2)^2 = 4$$

$$(ii) f(x) = \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}(x-2)^2}, \quad -\infty < x < \infty$$

$$\begin{aligned} &= \frac{1}{\sqrt{2} \times \sqrt{2} \sqrt{\pi}} e^{-\frac{1}{2 \times 2}(x-2)^2} \\ &= \frac{1}{\sqrt{2} \sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-2}{\sqrt{2}}\right)^2} \quad \dots(1), -\infty < x < \infty \end{aligned}$$

Comparing (1) with,

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

we get

$$\mu = 2, \quad \sigma = \sqrt{2}$$

$$\therefore \text{Mean} = \mu = 2 \text{ and variance} = \sigma^2 = (\sqrt{2})^2 = 2$$

$$3) \text{ i) } X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

$$\Rightarrow X_1 + X_2 \sim N(30 + 40, 25 + 35)$$

$$\Rightarrow X_1 + X_2 \sim N(70, 60)$$

$$\text{ii) } X_1 - X_2 \sim N(30 - 40, 25 + 35)$$

$$\Rightarrow X_1 - X_2 \sim N(-10, 60)$$

Check Your Progress 2

$$1) \text{ As } X \sim U[-a, a],$$

$\therefore$ probability density function of X is

$$f(x) = \frac{1}{a - (-a)} = \frac{1}{a + a} = \frac{1}{2a}, \quad -a < x < a.$$

$$\text{i) Given that } P[X > 4] = \frac{1}{3}$$

$$\Rightarrow \int_4^a \frac{1}{2a} dx = \frac{1}{3}$$

$$\Rightarrow \frac{1}{2a} [x]_4^a = \frac{1}{3}$$

$$\Rightarrow \frac{a - 4}{2a} = \frac{1}{3}$$

$$\Rightarrow 3a - 12 = 2a$$

$$\Rightarrow a = 12.$$

$$\text{ii) } P[X < 1] = \frac{3}{4}$$

$$\Rightarrow \int_{-a}^1 \frac{1}{2a} dx = \frac{3}{4}$$

$$\Rightarrow \frac{1}{2a} [x]_{-a}^1 = \frac{3}{4}$$

$$\Rightarrow \frac{1}{2a} [1 + a] = \frac{3}{4}$$

$$\Rightarrow 1 + a = \frac{3}{2} a$$

$$\Rightarrow 2 + 2a = 3a$$

$$\Rightarrow a = 2$$

$$2) \text{ As } X \sim U[-2, 2],$$

$$\therefore f(x) = \frac{1}{4}, \quad -2 < x < 2.$$

$$\text{Now } P[X > k] = \frac{1}{2}$$

$$\Rightarrow \int_k^2 \frac{1}{4} dx = \frac{1}{2}$$

$$\Rightarrow \frac{2-k}{4} = \frac{1}{2}$$

$$\Rightarrow 2 - k = 2$$

$$\Rightarrow k = 0.$$

Check Your Progress 3

- 1) We know that the mean and variance of chi-square distribution with n degrees of freedom are

$$\text{Mean} = n \text{ and Variance} = 2n$$

$$\text{In our case, } n = 10, \text{ therefore, Mean} = 10 \text{ and Variance} = 20.$$

- 2) Refer to Section 7.4.

Check Your Progress 4

- 1) Here, we are given that

$$\mu = 60, n = 15, \bar{X} = 65 \text{ and } S = 12$$

We know that the t-variate is

$$t = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

Therefore, we have

$$t = \frac{65 - 60}{12/\sqrt{15}} = \frac{5}{4.39} = 1.14$$

- 2) We know that the mean and variance of t-distribution with n degrees of freedom are

$$\text{Mean} = 0 \text{ and Variance} = \frac{n}{n-2} \text{ for } n > 2$$

$$\text{In our case, } n=20, \text{ therefore, Mean} = 0 \text{ \& Variance} = \frac{20}{20-2} = \frac{10}{9}$$

- 3) Refer to Section 7.5

Check Your Progress 5

- 1) Here, we are given that

$$n_1 = 15, n_2 = 10, \sigma_1 = 65, \sigma_2 = 45, S_1 = 60, S_2 = 50$$

The value of F-variate can be calculated as follows:

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$$

where, S_1^2 & S_2^2 are the values of sample variances.

Therefore, we have

$$F = \frac{(60)^2 / (65)^2}{(50)^2 / (45)^2}$$

$$= \frac{0.85}{1.23} = 0.69$$

- 2) Refer to Section 7.6.
- 3) We know that the mean and variance of F-distribution with v_1 and v_2 degrees of freedom are

$$\text{Mean} = \frac{v_2}{v_2 - 2} \quad \text{for } v_2 > 2 \quad \text{and}$$

$$\text{Variance} = \frac{2v_2^2(v_1 + v_2 - 2)}{v_1(v_2 - 2)^2(v_2 - 4)} \quad \text{for } v_2 > 4.$$

In our case, $v_1 = 5$ and $v_2 = 10$, therefore,

$$\text{Mean} = \frac{v_2}{v_2 - 2} = \frac{10}{10 - 2} = \frac{5}{4} \quad \text{and}$$

$$\text{Variance} = \frac{2(10)^2(5 + 10 - 2)}{5(10 - 2)^2(10 - 4)} = \frac{65}{48}$$

- 4) Refer to Section 7.6

7.9 REFERENCES

IGNOU Course Material

1. Post-Graduate Diploma in Applied Statistics (PGDAST), MST 003: Probability Theory, Block 4: Continuous Probability Distributions, UNIT 12: Normal Distribution (for Section 7.2 and related portions in Sections 7.0, 7.1, 7.7, 7.8 and Check Your Progress questions)
2. PGDAST, MST 003: Probability Theory, Block 4: Continuous Probability Distributions, UNIT 15: Continuous Uniform and Exponential Distributions (for Section 7.3 and related portions in Sections 7.0, 7.1, 7.7, 7.8 and Check Your Progress questions)
3. PGDAST, MST 004: Statistical Inference, Block 1: Sampling Distributions, UNIT 3: Standard Sampling Distributions - I (for Sections 7.4, 7.5 and related portions in Sections 7.0, 7.1, 7.7, 7.8 and Check Your Progress questions)
4. PGDAST, MST 004: Statistical Inference, Block 1: Sampling Distributions, UNIT 4: Standard Sampling Distributions - II (for Section 7.6 and related portions in Sections 7.0, 7.1, 7.7, 7.8 and Check Your Progress questions)

Reference Book

5. "Introduction to Probability and Statistics for Engineers and Scientists", 3rd Edition, Sheldon M. Ross, Elsevier Academic Press, 2004