UNIT 13 NORMAL DISTRIBUTION

Structure

13.1 INTRODUCTION

In Units 9 to 12, we have studied standard discrete distributions. From this unit onwards, we are going to discuss standard continuous univariate distributions. This unit and the next unit deal with normal distribution. Normal distribution has wide spread applications. It is being used in almost all data-based research in the field of agriculture, trade, business, industry and the 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 normal distribution has a unique position in probability theory, and it can be used as approximation to most of the other distributions. Discrete distributions occurring in practice including binomial, Poisson, hypergeometric, etc. already studied in the previous block (Block 3) can also be approximated by normal distribution. You will notice in the subsequent courses that theory of estimation of population parameters and testing of hypotheses on the basis of sample statistics have also been developed using the concept of normal distribution as most of the sampling distributions tend to normality for large samples. Therefore, study of normal distribution is very important.

Due to various properties and applications of the normal distribution, we have covered it in two units – Units 13 and 14. In the present unit, normal distribution is introduced and explained in Sec. 13.2. Chief characteristics of normal distribution are discussed in Sec. 13.3. Secs. 13.4, 13.5 and 13.6 describes the moments, mode, median and mean deviation about mean of the distribution.

Objectives

After studing this unit, you would be able to:

- introduce and explain the normal distribution;

- know the conditions under which binomial and Poisson distributions tend to normal distribution;
- state various characteristics of the normal distribution;
- compute the moments, mode, median and mean deviation about mean of the distribution; and
- solve various practical problems based on the above properties of normal distribution.

13.2 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:

i) n, the number of trials, is indefinitely large i.e. $n \to \infty$;

ii) 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 standardized 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.

For binomial distribution, you have already studied [see Unit 9 of this course] that

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\left[npq(q-p)\right]^2}{\left[npq\right]^3} = \frac{(q-p)^2}{npq},$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{npq\left[1+3(n-2)pq\right]}{\left[npq\right]^2} = 3 + \frac{1-6pq}{npq},$$

$$\gamma_1 = \sqrt{\beta_1} = \frac{q-p}{\sqrt{npq}} = \frac{1-2p}{\sqrt{npq}}, \text{ and}$$

$$\gamma_2 = \beta_2 - 3 = \frac{1-6pq}{npq}.$$

From the above results, it may be noticed that if $n \to \infty$, then moment coefficient of skewness $(\gamma_1) \to 0$ and the moment coefficient of kurtosis i.e. $\beta_2 \to 3$ or $\gamma_2 \to 0$. Hence, as $n \to \infty$, the distribution becomes symmetrical and the curve of the distribution becomes mesokurtic, which is the main feature of normal distribution.

Normal Distribution as a Limiting Case of Poisson Distribution

You have already studied in Unit 10 of this course that Poisson distribution is a limiting case of binomial distribution under the following conditions:

i) n, the number of trials is indefinitely large i.e. $n \to \infty$

ii) p, the constant probability of success for each trial is very small i.e. $p \to 0$.

iii)np is a finite quantity say 'λ'.

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

$Z = \dfrac{X-\lambda}{\sqrt{\lambda}}$ as λ increases indefinitely.

For Poisson distribution, you have already studied in Unit 10 of the course that

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{\lambda^2}{\lambda^3} = \frac{1}{\lambda} \Rightarrow \gamma_1 = \sqrt{\beta_1} = \frac{1}{\sqrt{\lambda}};\ \text{and}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\lambda^2}{\lambda^2} + \lambda = 3 + \frac{1}{\lambda} \Rightarrow \gamma_2 = \beta_2 - 3 = \frac{1}{\lambda}.$$

Like binomial distribution, here in case of Poisson distribution also it may be noticed from the above results that the moment coefficient of skewness $(\gamma_1) \to 0$ and the moment coefficient of kurtosis i.e. $\beta_2 \to 3$ or $\gamma_2 \to 0$ as $\lambda \to \infty$. Hence, as $\lambda \to \infty$, the distribution becomes symmetrical and the curve of the distribution becomes mesokurtic, which is the main feature of normal distribution.

Under the conditions discussed above, a random variable following a binomial distribution or following a Poisson distribution approaches to follow normal distribution, which 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(μ, σ^2)" and is expressed as $X \sim N(\mu, \sigma^2)$.

iii) No continuous probability function and hence 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 interval. For detail discussion on the concept, Sec. 5.4 of Unit 5 may be referred to.

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 (see Unit 8 of this course).

Mean of Z i.e. $E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X-\mu)\right]$

$$= \frac{1}{\sigma}\left[E(X)-\mu\right]$$

$$= \frac{1}{\sigma}\left[\mu-\mu\right] = 0 \qquad [\because E(X) = \text{Mean of } X = \mu\ ]$$

Variance of Z i.e. $V(Z) = V\left(\frac{X-\mu}{\sigma}\right)$

$$= \frac{1}{\sigma^2}\left[V(X-\mu)\right] = \frac{1}{\sigma^2}\left[V(X)\right]$$

$$= \frac{1}{\sigma^2}\left(\sigma^2\right) \qquad [\because \text{variance of X is } \sigma^2]$$

$$= 1.$$

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

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. 13.1):

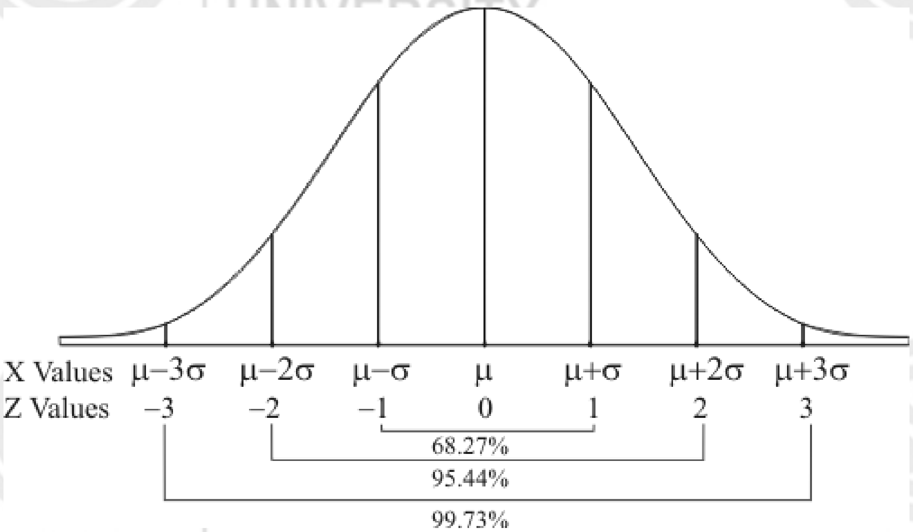

Fig. 13.1

Normal distribution has various properties and large number of applications. It can be used as 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 course MST-004) 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

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

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

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

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

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

(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

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

$$= \frac{1}{\sqrt{20}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-(-36)}{\sqrt{20}}\right)^2} = \frac{1}{\sqrt{40\pi}} e^{-\frac{1}{40}(x+36)^2}$$

$$= \frac{1}{2\sqrt{10\pi}} e^{-\frac{1}{40}(x+36)^2}, \qquad -\infty < x < \infty$$

(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

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} = \frac{1}{\sqrt{2}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{\sqrt{2}}\right)^2}$$

$$= \frac{1}{2\sqrt{\pi}} e^{-\frac{1}{4}x^2}, \qquad -\infty < x < \infty$$

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

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

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

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

Comparing it with,

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

we have

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

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

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

$= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{X-60}{4}\right)^2}$

Comparing it with,

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

we get

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

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

Here are some exercises for you.

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

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

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

E 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.

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

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

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

13.3 CHIEF 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. 13.1 given in Remark (vi) of Sec. 13.2.

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

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

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = 0,\ \beta_2 = \frac{\mu_4}{\mu_2^2} = 3$$

i.e. 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 – Median = 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) Q.D. : M.D. : S.D. = $\frac{2}{3}\sigma : \frac{4}{5}\sigma : \sigma$ = 10 : 12 : 15

xii) The points of inflexion of the curve are

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

xiii) If $X_1, X_2, ..., X_n$ are independent normal variables with means $\mu_1, \mu_2, ..., \mu_n$ and variances $\sigma_1^2, \sigma_2^2, ..., \sigma_n^2$ respectively then the linear combination $a_1X_1 + a_2X_2 + ... + a_nX_n$ of $X_1, X_2, ..., X_n$ is also a normal variable with mean $a_1\mu_1 + a_2\mu_2 + ... + a_n\mu_n$ and variance $a_1^2\sigma_1^2 + a_2^2\sigma_2^2 + ... + a_n^2\sigma_n^2$.

xiv) 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

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

Also, if $X_1, X_2, \ldots, X_n$ are independent variates each distributed as $N(\mu, \sigma^2)$, then their mean $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$.

xv) Area property:

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

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

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

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

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

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

This property and its applications will be discussed in detail in Unit 14.

Let us now establish some of these properties.

13.4 MOMENTS OF NORMAL DISTRIBUTION

Before finding the moments, following is defined as gamma function [See Unit 16 of the Course also for detail discussion] which is used for computing the even order central moments.

Gamma Function

If n > 0, the integral $\int_0^{\infty} x^{n-1}e^{-x}dx$ is called a gamma function and is denoted by $\Gamma(n)$.

$$\text{e.g.} \quad \int_0^{\infty} x^2 e^{-x}dx = \int_0^{\infty} x^{3-1}e^{-x}dx = \Gamma(3)$$

$$\text{and} \quad \int_0^{\infty} x^{-1/2} e^{-x}dx = \int_0^{\infty} x^{\frac{1}{2}-1}e^{-x}dx = \Gamma\left(\frac{1}{2}\right)$$

Some properties of the gamma function are

i) If n > 1, $\Gamma(n) = (n-1)\Gamma(n-1)$

ii) If n is a positive integer, then $\Gamma n = \underline{|n-1}$

iii) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

Now, the first four central moments of normal distribution are obtained as follows:

First Order Central Moment

As first order central moment (μ_1) of any distribution is always zero [see Unit 3 of MST-002], therefore, first order central moment (μ_1) of normal distribution = 0.

Second Order Central Moment

$\mu_2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$ [See Unit 8 of MST-003]

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

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x-\mu = \sigma z$

Differentiating

$\frac{dx}{\sigma} = dz$

$\Rightarrow dx = \sigma dz$

Also, when $x \to -\infty$, we have $z \to -\infty$ and

and when $x \to \infty$, $z \to \infty$

$$\therefore \ \mu_2 = \int_{-\infty}^{\infty} (\sigma z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^2 e^{-\frac{1}{2}z^2} dz$$

$$= \frac{2\sigma^2}{\sqrt{2\pi}} \int_{0}^{\infty} z^2 e^{-\frac{z^2}{2}} dz$$

$\because$ on changing z to – z, the integrand i.e. $z^2 e^{-\frac{z^2}{2}}$ does not get changed i.e. it is an even function of z [see Unit 2 of MST-001]. Now, the following property of definite integral can be used:

$\int_{-\infty}^{\infty} f(z)dz = 2\int_{0}^{\infty} f(z)dz$ if $f(z)$ is even function of z

Now, put $z^2 = 2t \Rightarrow z = \sqrt{2}\sqrt{t} = \sqrt{2}\ t^{\frac{1}{2}} \Rightarrow dz = \sqrt{2}\frac{1}{2}t^{-\frac{1}{2}}dt$

$$\therefore\ \mu_2 = \sigma^2\sqrt{\frac{2}{\pi}}\int_0^{\infty}(2t)e^{-t}\frac{1}{\sqrt{2}t^{\frac{1}{2}}}dt = \sigma^2\sqrt{\frac{2}{\pi}}\sqrt{2}\int_0^{\infty}t^{\frac{1}{2}}e^{-t}dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\int_0^{\infty}t^{\frac{3}{2}-1}e^{-t}dt$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\left\lceil\left(\frac{3}{2}\right)\right. \qquad \text{[By def. of gamma function]}$$

$$= \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\left\lceil\frac{1}{2}\right. \qquad \text{[By Property (i) of gamma function]}$$

$$= \frac{\sigma^2}{\sqrt{\pi}}\left(\sqrt{\pi}\right) \qquad \text{[By Property (iii) of gamma function]}$$

$$= \sigma^2$$

Third Order Central Moment

$$\mu_3 = \int_{-\infty}^{\infty}(x-\mu)^3 f(x)dx$$

$$= \int_{-\infty}^{\infty}(x-\mu)^3\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dx$$

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x-\mu = \sigma z \Rightarrow dx = \sigma dz$

and hence

$$\mu_3 = \int_{-\infty}^{\infty}(\sigma z)^3 . \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}z^2}\sigma dz$$

$$= \sigma^3\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^3 e^{-\frac{1}{2}z^2}dz$$

Now, as integrand $z^3 e^{-\frac{1}{2}z^2}$ changes to $-z^3 e^{-\frac{1}{2}z^2}$ on changing z to – z i.e. $z^3 e^{-\frac{1}{2}z^2}$ is an odd function of z.

Therefore, using the following property of definite integral:

$$\boxed{\int_{-a}^{a} f(z)dz = 0 \text{ if f(z) is an odd function of z}}$$

we have,

$$\mu_3 = \sigma^3\frac{1}{\sqrt{2\pi}}(0) = 0$$

Fourth Order Central Moment

$$\mu_4 = \int_{-\infty}^{\infty} (x-\mu)^4 f(x)\,dx = \int_{-\infty}^{\infty} (x-\mu)^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

Putting $\frac{x-\mu}{\sigma} = z$

$\Rightarrow$ $dx = \sigma dz$

$$\therefore \mu_4 = \int_{-\infty}^{\infty} (\sigma z)^4 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma^4}{\sqrt{2\pi}} \int_{-\infty}^{\infty} z^4 e^{-\frac{1}{2}z^2} dz = \frac{2\sigma^4}{\sqrt{2\pi}} \int_{0}^{\infty} z^4 e^{-\frac{1}{2}z^2} dz$$

$\because$ integrand $z^4.e^{-\frac{z^2}{2}}$ does not get changed on changing z to – z and hence it is an even function of z [using the same property as used in case of μ_2].

Put $\frac{z^2}{2} = t \Rightarrow z^2 = 2t$

$\Rightarrow 2zdz = 2dt$

$\Rightarrow$ $z\,dz = dt$

$\Rightarrow dz = \frac{dt}{z} = \frac{dt}{\sqrt{2t}}$

$$\therefore \mu_4 = \frac{2\sigma^4}{\sqrt{2\pi}} \int_{0}^{\infty} (2t)^2 e^{-t} \frac{1}{\sqrt{2t}} dt$$

$$= \frac{2\sigma^4.4}{\sqrt{2\pi}\sqrt{2}} \int_{0}^{\infty} t^2 e^{-t} \frac{1}{\sqrt{t}} dt = \frac{4\sigma^4}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{3}{2}} e^{-t} dt$$

$$= \frac{4\sigma^4}{\sqrt{\pi}} \int_{0}^{\infty} t^{\frac{5}{2}-1} e^{-t} dt$$

$= \frac{4\sigma^4}{\sqrt{\pi}} \left|\frac{5}{2}\right.$ [By definition of gamma function]

$= \frac{4\sigma^4}{\sqrt{\pi}} \frac{3}{2} \left|\frac{3}{2}\right.$ [By Property (i) of gamma function]

$= \frac{4\sigma^4}{\sqrt{\pi}} \frac{3}{2} \frac{1}{2} \left|\frac{1}{2}\right.$ [By Property (i) of gamma function]

$= \frac{3\sigma^4}{\sqrt{\pi}} \sqrt{\pi}$ [Using $\left|\frac{1}{2}\right. = \sqrt{\pi}$ (Property (iii) of gamma function)]

$= 3\sigma^4$

Thus, the first four central moments of normal distribution are

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

$$\Rightarrow \beta_1 = \frac{\mu_3^{\ 2}}{\mu_2^{\ 3}} = 0,\ \beta_2 = \frac{\mu_4}{\mu_2^{\ 2}} = \frac{3\sigma^4}{\left(\sigma^2\right)^2} = \frac{3\sigma^4}{\sigma^4} = 3$$

Therefore, moment coefficient of skewness $(\gamma_1) = 0$

$\Rightarrow$ the distribution is symmetrical.

The moment coefficient of kurtosis is $\beta_2 = 3$ or $\gamma_2 = 0.$

$\Rightarrow$ The curve of the normal distribution is mesokurtic.

Now, let us obtain the mode and median for normal distribution in the next section.

13.5 MODE AND MEDIAN OF NORMAL DISTRIBUTION

Mode

Let $X \sim N\left(\mu, \sigma^2\right)$, then p.d.f. of X is given by

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

$$, -\infty < x < \infty$$

Taking logarithm on both sides of (1), we get

$$\log f(x) = \log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2 \log e \qquad \left[\begin{array}{l}\because \log mn = \log m + \log n \\ \text{and} \log m^n = n \log m\end{array}\right]$$

$$= \log\frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2\sigma^2}(x-\mu)^2 \qquad \text{as } \log e = 1$$

Differentiating w.r.t.x

$$\frac{1}{f(x)} f'(x) = 0 - \frac{1}{2\sigma^2} 2(x-\mu) = -\frac{(x-\mu)}{\sigma^2}$$

$$\Rightarrow f'(x) = -\frac{(x-\mu)}{\sigma^2} f(x) \qquad \ldots (2)$$

For maximum or minimum

$f'(x) = 0$

$$\Rightarrow -\frac{(x-\mu)}{\sigma^2} f(x) = 0$$

$$\Rightarrow x - \mu = 0 \text{ as } f(x) \neq 0$$

$\Rightarrow x = \mu$

Now differentiating (2) w.r.t. x, we have

$$f''(x) = -\frac{(x-\mu)}{\sigma} f'(x) - \frac{(1)}{\sigma^2} f(x)$$

$$f''(x)\Big|_{\text{at } x=\mu} = 0 - \frac{f(\mu)}{\sigma^2} = -\frac{f(\mu)}{\sigma^2} < 0$$

$\therefore x = \mu$ is point where function has a maximum value.

$\Rightarrow$ Mode of X is μ.

Median

Let M denote the median of the normally distributed random variable X.

We know that median divides the distribution into two equal parts

$$\therefore \int_{-\infty}^{M} f(x)dx = \int_{M}^{\infty} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{\mu} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

In the first integral, let us put $\frac{x-\mu}{\sigma} = z$

Therefore, $dx = \sigma\, dz$

Also when $x = \mu \Rightarrow z = 0$, and

when $x \to -\infty \Rightarrow z \to -\infty$.

Thus, we have

$$\int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z)^2} dz + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} dz + \int_{\mu}^{M} f(x)dx = \frac{1}{2}$$

$$\Rightarrow \frac{1}{2} + \int_{\mu}^{M} f(x)\,dx = \frac{1}{2} \qquad \left[\begin{array}{l} \because Z \text{ is s.n.v. with p.d.f. } \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \\ \text{So } \int_{-\infty}^{\infty} \phi(z)dz = 1 \Rightarrow \int_{-\infty}^{0} \phi(z)dz = \frac{1}{2} \end{array}\right]$$

$$\Rightarrow \int_{\mu}^{M} f(x)dx = 0$$

$$\Rightarrow M = \mu \qquad \text{as } f(x) \neq 0$$

$\therefore$ Median of $X = \mu$

From the above two results, we see that

$\boxed{\textbf{Mean = Median = Mode = } \mu}$

13.6 MEAN DEVIATION ABOUT THE MEAN

Mean deviation about mean for normal distribution is

$$= \int_{-\infty}^{\infty} |x - \text{Mean}|\, f(x) dx \qquad \text{[See Section 8.4 of Unit 8]}$$

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

Put $\frac{x-\mu}{\sigma} = z \Rightarrow x - \mu = \sigma z$

$\Rightarrow \frac{dx}{\sigma} = dz$

$$\therefore \text{ M.D. about mean} = \int_{-\infty}^{\infty} |\sigma z| \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \sigma dz$$

$$= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |z| e^{-\frac{1}{2}z^2} dz$$

Now, $|z| e^{-\frac{1}{2}z^2}$ (integrand) is an even function z as it does not get changed on changing z to –z, $\therefore$ by the property,

" $\int_{-a}^{a} f(x)dx = 2\int_{0}^{a} f(x)dx$, if $f(x)$ is an even function of x", we have

$$\text{M.D. about mean} = \frac{\sigma}{\sqrt{2\pi}} 2\int_{0}^{\infty} |z| e^{-\frac{1}{2}z^2} dz$$

Now, as the range of z is from 0 to ∞ i.e. z takes non-negative values,

$\therefore$ $|z|$ = z and hence

$$\text{M.D. about mean} = \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} z e^{-\frac{1}{2}z^2} dz$$

Put $\frac{z^2}{2} = t \Rightarrow z^2 = 2t \Rightarrow 2zdz = 2dt \Rightarrow zdz = dt$

$$\therefore \text{ M.D. about mean} = \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-t} dt = \sqrt{2}\frac{\sigma}{\sqrt{\pi}} \left[\frac{e^{-t}}{-1}\right]_0^{\infty} = \sqrt{\frac{2}{\pi}}\sigma[-0+1] = \sqrt{\frac{2}{\pi}}\sigma$$

In practice, instead of $\sqrt{\frac{2}{\pi}}\sigma,$ its approximate value is mostly used and that is $\frac{4}{5}\sigma.$

$$\because \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2\times 7}{22}} = \sqrt{\frac{7}{11}} = \sqrt{0.6364} = 0.7977 = 0.08 \text{ or } \frac{4}{5}\text{(approx.)}$$

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

13.7 SOME PROBLEMS BASED ON PROPERTIES OF NORMAL DISTRIBUTION

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.

$X_1 \sim N\left(\mu_1, \sigma_1^2\right)$ and $X_2 \sim N\left(\mu_2, \sigma_2^2\right)$

then

$X_1 + X_2 \sim N\left(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\right)$, and

$X_1 - X_2 \sim N\left(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2\right)$ [See Property xiii (Section 13.3)]

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

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

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

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

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

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

Example 5: If $X \sim N(0, 1)$, what are its first four central moments?

Solution: Here $\mu = 0$, $\sigma^2 = 1 \Rightarrow \sigma = 1$.

$\therefore$ first four central moments are:

$\mu_1 = 0,\ \mu_2 = \sigma^2 = 1,\ \mu_3 = 0,\ \mu_4 = 3\sigma^4 = 3.$

Example 6: If X_1, X_2 are independent variates such that $X_1 \sim N(40, 25)$, $X_2 \sim N(60, 36)$, then find mean and variance of (i) $X = 2X_1 + 3X_2$ (ii) $Y = 3X_1 - 2X_2$

Solution: Here $X_1 \sim N(40, 25)$, $X_2 \sim N(60, 36)$

$\therefore$ Mean of $X_1 = E(X_1) = 40$

Variance of $X_1 = Var(X_1) = 25$

Mean of $X_2 = E(X_2) = 60$

Variance of $X_2 = Var(X_2) = 36$

Now,

(i) Mean of X = $E(X) = E(2X_1 + 3X_2) = E(2X_1) + E(3X_2)$

$$= 2E(X_1) + 3E(X_2) = 2\times 40 + 3\times 60 = 80 + 180 = 260$$

$$Var(X) = Var(2X_1 + 3X_2)$$

$$= Var(2X_1) + Var(3X_2) \qquad [\because X_1 \text{ and } X_2 \text{ are independent}]$$

$$= 4Var(X_1) + 9Var(X_2)$$

$$= 4\times 25 + 9\times 36 = 100 + 324 = 424$$

(ii) Mean of Y = $E(Y) = E(3X_1 - 2X_2)$

$$= E(3X_1) + E(-2X_2)$$

$$= 3E(X_1) + (-2)E(X_2)$$

$$= 3\times 40 - 2\times(60) = 120 - 120 = 0$$

$$Var(Y) = Var(3X_1 - 2X_2)$$

$$= Var(3X_1) + Var(-2X_2)$$

$$= (3)^2 Var(X_1) + (-2)^2 Var(X_2)$$

$$= 9\times 25 + 4\times 36 = 225 + 144 = 369$$

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

E3) 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

i) $X_1 + X_2$

ii) $X_1 - X_2$

E4) If X ~ N(50, 225), find its Quartile deviation.

E5) If X_1 and X_2 are independent variates with each distributed as N (50, 64), what is the distribution of $\frac{X_1 + X_2}{2}$?

E6) For a normal distribution, the first moment about 5 is 30 and the fourth moment about 35 is 768. Find the mean and standard deviation of the distribution.

13.8 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 σ^2(>0) if it takes on any real value and its probability density function is given by

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

2) If $X \sim N\left(\mu,\sigma^2\right)$, then $Z=\dfrac{X-\mu}{\sigma}$ 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) Q_3 – Median = Median – Q_1

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

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

8) Central moments of Normal distribution are

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

9) Moment coefficient of skewness is zero and the curve is always mesokurtic.

10) Sum of independent normal variables is also a normal variable.

13.9 SOLUTIONS/ANSWERS

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

$\therefore$ in usual notations, we have

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

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

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

$$=\frac{1}{\frac{2}{3}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-1/2}{2/3}\right)^2}$$

$$= \frac{3}{2\sqrt{2\pi}} e^{-\frac{9}{2}\left(\frac{2x-1}{4}\right)^2}, \qquad -\infty < x < \infty$$

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

$\therefore$ in usual notations, we have

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

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

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

$$= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-(-40)}{4}\right)^2}$$

$$= \frac{1}{4\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x+40}{4}\right)^2}, \quad -\infty < x < \infty$$

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

$$= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\frac{x^2}{4}}$$

$$= \frac{1}{2\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-0}{2}\right)^2} \qquad \ldots(1) \qquad , -\infty < x < \infty$$

Comparing (1) with,

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

we get

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

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

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

$$= \frac{1}{\sqrt{2}\times\sqrt{2}\sqrt{\pi}} e^{-\frac{1}{2\times 2}(x-2)^2}$$

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

Comparing (1) with,

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

we get

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

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

E3) i) $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

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

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

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

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

E4) As $\sigma^2 = 225$

$\Rightarrow \sigma = 15$

and hence Q.D. $= \frac{2}{3}\sigma = \frac{2}{3} \times 15 = 10$

E5) We know that if $X_1, X_2, .., X_n$ are independent variates each distributed as $N(\mu, \sigma^2)$, then $\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$

Here X_1 and X_2 are independent variates each distributed as N(50, 64),

$\therefore$ their mean i.e. $\bar{X}$ i.e. $\frac{X_1 + X_2}{2} \sim N\left(50, \frac{64}{2}\right)$

i.e $\bar{X} \sim N(50, 32)$.

E6) We know that $\mu_1' = \bar{x} - A$ [See Unit 3 of MST-002]

where μ_1' is the first moment about A.

$\therefore 30 = \bar{x} - 5 \Rightarrow \bar{x} = 35 \Rightarrow$ Mean = 35

Given that fourth moment about 35 is 768. But mean is 35, and hence the fourth moment about mean = 768.

$\Rightarrow \mu_4 = 768$

$\Rightarrow 3\sigma^4 = 768$

$\Rightarrow \sigma^4 = \frac{768}{3}$ $\left[\because \mu_4 = 3\sigma^4\right]$

$\Rightarrow \sigma^4 = 256 = (4)^4 \Rightarrow \sigma = 4.$