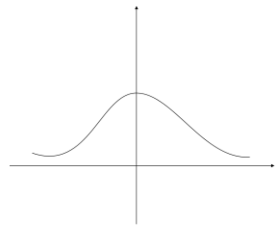
NORMAL DISTRIBUTION



Normal

Normal distribution is a continuous distribution. The normal distribution is a limiting case if Binomial probability distribution if

(i) The number if trials or the value of n is very large (nearly infinity).
(ii) If neither p or q is very small.

If p and q are nearly equal, the normal distribution is very close to Binomial distribution even though the value of n is **not** large.

The normal distribution can be regarded as a limiting case of the Poisson distribution if the value of the mean i.e. $m \text{ or } \lambda \text{ or } \mu$ is very large nearly infinity.

Importance of normal distribution

The normal distribution is often called the corner stone of statistical analysis and drawing of inferences.

1. If a random sample is taken from any universe, then as the sample size increases the mean of the sample approaches the normal distribution (with mean μ and standard deviation $\sigma/\sqrt{n}$). This is the **central limit theorem.** This property of the normal distribution is by far the most important as it enables us to draw inferences about the universe by making sample studies. According to **central limit theorem** we can estimate the upper and lower limits within which a value in the universe lie, by conducting sample studies.
2. Even though the assumptions of a normal distribution are not satisfied, the results given by a normal distribution study, in many cases, are found to be highly satisfactory. However, theoretically the results hold good under the assumptions that the properties of the normal distribution are applicable to the problem under study.
3. There are many mathematical properties possessed by this distribution which makes its application possible in a wide variety of situations and for making varied types of studies.
4. It is very useful in statistical quality control where the control limits are set by using this distribution.

The normal distribution is given by the equation

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

Where $\mu = mean$, $\sigma =$ standard deviation, $\pi = 3.14159........$, $e = 2.71828.......$

$$P(x_1 < x < x_2) = \int_{x_1}^{x_2} \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

On substituting $z = \frac{x-\mu}{\sigma}$ in (1), we get

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-1}{2}z^2} \text{....................................} (2)$$

Hence mean $= 0$ and standard deviation $= 1$.

(2) is known as standard form of the normal distribution.

Properties of the normal distribution.

1. The normal curve is symmetrical about the mean, i.e. there is no skewness in it. The number of cases above the mean value and below the mean value are equal.
2. The mean, median and mode have the same value i.e. they coincide. Thus in a normal distribution or curve

   Mean=Medan=Mode
3. The height of the normal is maximum at the mean value. This ordinate divides the curve in two equal parts or identical parts.
4. Since the curve is symmetrical, the first and the third quartile are equidistant from the median. Thus $Q_3 - Median = Median - Q_1$
5. Since there is only one point of maximum frequency (at the mean) the normal distribution is uni-modal.
6. The curve is asymptotic to the baseline. It continues to approach but never touches the baseline.

The standard normal curve

A normal distribution in which the mean is **zero** and the standard deviation is **unity** is known as the standard normal curve.

Its utility is very great because curves with mean $\overline{X}$ and standard deviation σ can be converted into a standard normal curve by change of origin and scale. This becomes necessary as otherwise in different distributions with different values of mean and standard deviation, it would be very difficult to find out the area between various ordinates. Once they are converted into standard normal curve this problem is solved.

In the original scale the mean and standard deviation are $\overline{X}$ and σ, but in the new scale which is called the $z - scale$ the mean is **zero** and the standard deviation is **unity .** This process of changing the $X - scale$ is called $z - transformation$. Z is called the standard normal variate and is given by

$$Z = \frac{X - \overline{X}}{\sigma}$$

Where $\overline{X}$ is the mean and σ the standard deviation of the given normal distribution , X is the observed value at which we want to find the value of z .

The values of z for different values of X, define a normal distribution with mean 0 and standard deviation is equal to unity. This new distribution is called **Standard normal distribution** or **unit normal distribution.** The curve drawn with this distribution is called **standard normal curve.**

Properties of standard normal distribution or curve.

1. The area under the standard normal curve is unity.
2. The mean of the standard normal curve (or distribution) $= 0$.
3. Standard deviation of standard normal distribution is equal to unity.

Any normal distribution regardless of the value of $\overline{X}$ or σ can be changed to standard normal distribution by defining standard normal variate

$$Z = \frac{X - \overline{X}}{\sigma}$$

Example 1

For a standardized random variable $Z \sim N(0.1)$.

Determine:

(i) $P(0 < z < 0.4)$

(ii) $P(1 < z < 2)$

(iii) $P(0.25 < z < 1.96)$

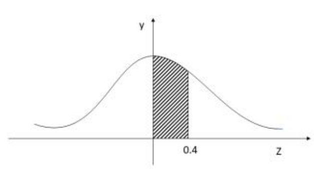
(iv) $P(z < -1.80)$

(v) $P(-1.75 < z < -0.83)$

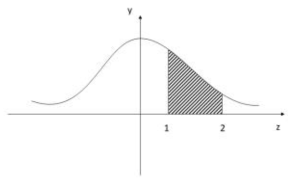
(vi) $P(-1.60 < z < 1.65)$

Solution

(i) $P(0 < z < 0.4) = \Phi(0.4) - \Phi(0) = 0.1554 - 0.000 = 0.1554$

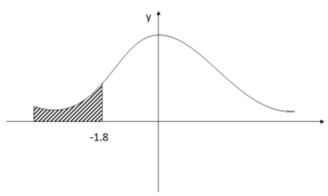


(ii) $P(1 < z < 2) = \Phi(2) - \Phi(1) = 0.4772 - 0.3413 = 0.1359$

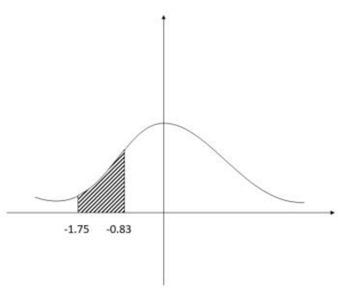


(iii) $P(0.25 < z < 1.96) = \Phi(1.96) - \Phi(0.25) = 0.4750 - 0.0987 = 0.3763$

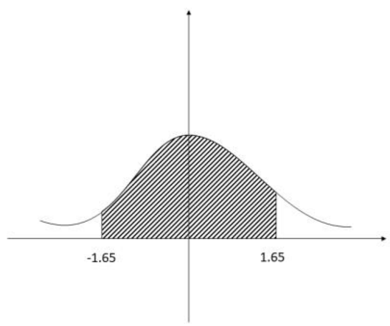
(iv) $P(z < -1.80) = 0.5 - \Phi(1.80) = 0.5 - 0.4641 = 0.0359$



(v) $P(-1.75 < z < -0.83) = \Phi(1.75) - \Phi(0.83) = 0.4599 - 0.2967 = 0.1632$



(vi) $P(-1.65 < z < 1.60) = 0.4505 + 0.4452 = 0.8957$



Example 2

Given a probability , what is the value of z ?

(i) Find c, if $P(0 \le z \le c) = 0.403$

(ii) Find d, if $P(z \le d) = 0.726$

(iii) Find e, if $P(-0.2 < z < e) = 0.266$

(iv) Find f, if $P(|z| \le f) = 0.980$

Solution

(i) $c = 1.3$

(ii) $\Phi(-d) = 0.726 - 0.5 = 0.226$

$-d = 0.6$

$d = -0.6$

(iii) $\Phi(-e) - \Phi(-0.2) = 0.266$

$\Phi(-e) - 0.0793 = 0.266$

$\Phi(-e) = 0.3453$

$-e = 1.2$

$e = -1.2$

(iv) $P(-f \le z \le f) = 0.980$

Remaining area $1 - 0.980 = 0.020$

Thus each part $= 0.01$

$\Phi(f) = 0.5 - 0.01 = 0.49$

$f = 2.31 \ or \ 2.32 \ or \ 2.34$

Example 3

In an intelligence test administered on 1000 students, the average was 42 and standard deviation 24. Find

(i) The number of students exceeding a score 50.
(ii) The number of students lying between 30 and 54.
(iii) The value of score exceeded by top 100 students.

Solution

(i) When score is 50, then

$$z = \frac{x-\bar{x}}{\sigma} = \frac{50-42}{24} = \frac{8}{24} = 0.33$$

$$P(z > 0.33) = 0.5 - \Phi(0.33) = 0.5 - 0.1293 = 0.3707$$

Number of students exceeding a score $50 = 0.3707 \times 1000 = 370.7 \approx 371$

(ii) When the score is 30, $z = \frac{30-42}{24} = -0.5$

When the score is 54, $z = \frac{54-42}{24} = 0.5$

$$P(-0.5 < z < 0.5) = \Phi(0.5) + \Phi(0.5) = 0.1915 + 0.1915 = 0.3830$$

$\therefore$ The number of students lying between a score 30 and $54 = 0.3830 \times 1000 = 383$

(iii) Area covered by top 100 students

$$\frac{100}{1000} = 0.1$$

Remaining area on the right of mean $= 0.5 - 0.1 = 0.4$ and the corresponding value of $z = 1.28$

x is the value of the score exceeded by top 100 students.

$$z = \frac{x-\bar{x}}{\sigma}$$

$$1.28 = \frac{x-42}{24}$$

$$\therefore x = 72.72 \approx 73$$

Example 4

In a certain examination, 31% of the students got less than 45 marks and 8% of the students got more than 64 marks. Assuming the distribution to be normal, find the mean and standard deviation of the marks.

Example 5

In a distribution exactly normal, 7% of the items are under 35 and 89% are under 63, what are mean and standard deviation of the distribution?

Example 6

Glass rods produced in a certain workshop have masses that follow the normal distribution with mean μ and standard deviation σ .Given that 8.85% of the rods have masses less than 20 grams and 4.27% have masses above 80 grams, determine the:

i. Values of μ and σ.
ii. Probability that a rod picked at random will have a mass less than 60 grams.

Example 7

The lifetimes of a certain type of a computer component manufactured by a certain factory are normally distributed with mean of 800 hours and a standard deviation of 160 hours.

(i) The manufacturer replaces all components that fail before the guaranteed minimum lifetime of 600 hours. Determine the percentage of the components that have to be replaced.

(ii) Determine the guaranteed lifetime if the manufacturer only wishes to replace 1% of the components.

(iii) Determine the probability that the mean lifetime of a sample of these computer components exceed 850 hours.

Example 8

The heights of students in a certain institution are normally distributed with a mean of 160 cm and a standard deviation of 10 cm .Determine the:

i. Probability that a student chosen at random will have a height greater than $175cm$;
ii. Proportion of students with heights between 143 cm and 168 cm;
iii. Number of students likely to have a height less than 152 cm, if the total number of students in the institution is 1000.

Example 9

The mean of a normal distribution is 50 and standard deviation is 12.5. Calculate the 60^{th}, 65^{th} and 90^{th} percentiles.

Example 10

The following tables gives frequencies of occurrence of a variate X between certain limits.

Variate X	Frequency
Less than 40	30
40 or more but less than 50	33
50 and more	37
Total	100

The distribution is exactly normal. Find the average and standard deviation.

Example 11

Assuming that the height distribution of a group of men is normal, find the mean and standard deviation, given that 84% of the men have heights less than 65.2 inches and 68% have height between 65.2 inches and 62.8 inches.

Example 12

Sacks of grain packed by an automatic machine loader have an average weight of 114 pounds. It is found that 11% of the bags are over 116 pounds. If the distribution is normal, find the standard deviation.

Example 13

Records kept by the goods inwards department of a large factory shows that the average number of lorries arriving each week is 248. It is known that the distribution approximates to the normal with a standard deviation of 26. If this pattern of arrival continues what percentage of weeks can be expected to have number of arrivals of

(i) Less than 229 per week?
(ii) More than 280 per week?

Example 14

Mean annual sales of a firm is worth £150 with a standard deviation of £20. For how many days in a year of 365 days his sales are expected to be worth less than £100.

.