CHAPTER V : Normal Distribution

Normal Distribution (most used statistical distribution)

⇒ Normally arises naturally in many physical, biological and social measurement situations.

⇒ Normally is important in statistical inference.

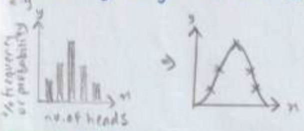
Characteristics:

- Bell shaped & is symmetrical about its mean.
- Asymptotic to axis, i.e., it extends indefinitely in either direction from the mean.
- continuous distribution.
- Family of curves, i.e., every unique pair of mean and standard deviation defines a different normal distribution. Thus, the normal distribution is completely described by 2 parameters: mean and S.D.
- Total area under the curve sums to 1, i.e., the area of distribution on each side of the mean is 0.5.
- It is unimodal, i.e., values mound up only in the center of the curve.
- The probability that a random variable will have a value between any two points is equal to the area under the curve between those points.

The Normal Curve

As the number of columns increases, the entire shape of the histogram begins to approximate a curve, with the shaded areas all under the top line. And, in fact, we can readily convert these histograms (using rectangles) to a curve by joining up the central points on top of each column.

The most frequent result is in the centre and the frequencies decline as one moves away from the centre (indicated by the decreasing height of the columns).

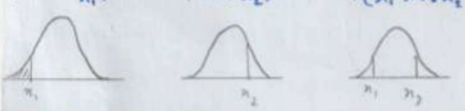


Finding Probabilities

The probability that X lies between a & b is written [illegible]. To find this probability, you need to find the [illegible] under the normal curve between a and b. One way [illegible] area by integrating the function but some [illegible] to integrate [illegible] tables are used instead.

Probabilities are given by areas under the normal curve.

$P(X < x_1)$ $P(X > x_2)$ $P(x_1 < X < x_2)$



Note that $P(X = 0) = 0$. This means that $P(X < x_1)$ is the same as $P(X \le x_1)$. $P(X > x_2)$ is the same as $P(X \ge x_2)$ and $P(x_1 < X < x_2)$ is the same as $P(x_1 \le X \le x_2)$

The Standard Normal Variable, Z

To use same set of tables for all possible of μ and σ^2, the variable X is standardised so that the mean is 0 and S.D is 1. Notice that since the variance is the square of the S.D, the variance is also 1. The standardised normal variable is called Z and $Z \sim N(0,1)$

In general:

To standardise X, where $X \sim N(\mu, \sigma^2)$

- subtract the mean, μ
- Then divide by the standard deviation, σ

To obtain:

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

Using standard Normal Tables:

The standard normal tables give the area under the curve as far as a particular value z. This is written as $\Phi(z)$

This area gives the probability that Z is less than a particular value z, so $P(Z < z) = \Phi(z)$

In general, $P(Z > a) = 1 - \Phi(a)$ ⇒ $P(Z > a) = 1 - P(Z < a)$

$P(Z > -a) = \Phi(a) = P(Z < a)$

$P(Z < -a) = 1 - \Phi(a) = P(Z > a)$

Using Standard Normal Tables For Any Normal Variable X

Remember that to standardize X, where $X \sim N(\mu, \sigma^2)$

- subtract the mean, μ
- Then divide by S.D, σ

To give $Z = \frac{X - \mu}{\sigma}$, where $Z \sim N(0,1)$

Example: Lengths of metal strips produced by a machine are normally distributed with mean length of 150 cm and a S.D of 10 cm. (keywords) Find probability that the length of a randomly selected strip is:

X - Length of metal strips

$X \sim N(150, 10^2)$

(a) shorter than 165 cm

$P(X < 165) = P(Z < \frac{165 - 150}{10})$

$= P(Z < 1.5)$

$= 0.9332$

(b) longer than 150 cm

$P(X > 150) = P(Z > \frac{150 - 150}{10})$

$= P(Z > 0)$

$= 1 - \Phi(0)$

$= 1 - 0.5000$

$= 0.5000$

Using The Standard Normal Tables In Reverse To Find Z when $\Phi(z)$ is known

To find z when you know $\Phi(z)$ you have to read the tables 'in reverse'. The following notation is useful.

If $\Phi(z) = p$ then $z = \Phi^{-1}(p)$

Example: If $Z \sim N(0,1)$, find the value of a if:

(a) $P(Z < a) = 0.9693$

$\Phi(a) = 0.9693$

$\therefore a = \Phi^{-1}(0.9693)$

$= 1.87$

(b) $P(Z > a) = 0.3802$

$1 - \Phi(a) = 0.3802$

$\Phi(a) = 0.6198$

$\therefore a = \Phi^{-1}(0.6198)$

$= 0.305$

(c) $P(Z > a) = 0.7367$

$P(Z < -a) = 0.7367$

$\Phi(-a) = 0.7367$

$-a = \Phi^{-1}(0.7367)$

$\therefore a = -0.633$

Using Tables in Reverse For Any Normal Variable X.

Example: The heights of female students [illegible] particular college are normally distributed with a mean of 16[illegible] cm and a S.D of 9 cm.

X - Height of th[illegible]

$X \sim N($[illegible]

(a) [illegible] that 80% of these female students have a height less than h cm, find the value [illegible].

$P(X < h) = 0.80$

$P(Z < \frac{h - 169}{9}) = 0.80$

$\Phi(\frac{h - 169}{9}) = 0.80$

$\frac{h - 169}{9} = \Phi^{-1}(0.80)$

$\therefore h = 176.578$

$= 177$ cm (3 s.f)

(b) Given that 60% of these female students have a greater height than s cm, find the value of s

$P(X > s) = 0.60$

$P(Z > \frac{s - 169}{9}) = 0.60$

$P(Z < \frac{-s + 169}{9}) = 0.60$

$\frac{-s + 169}{9} = \Phi^{-1}(0.60)$

$\frac{-s + 169}{9} = 0.253$

$\therefore s = 166.723$

Finding the value of μ or σ or both

The lengths of certain items follow a normal distribution with mean μ cm and S.D 6 cm. It is known that 4.78% of items have value greater than 82 cm. Find μ.

$P(X > 82) = 0.0478$ ← $4.78\% = \frac{4.78}{100} = 0.0478$

$P(Z > \frac{82 - \mu}{6}) = 0.0478$

$1 - \Phi(\frac{82 - \mu}{6}) = 0.0478$

$\Phi(\frac{82 - \mu}{6}) = 0.9522$

$\frac{82 - \mu}{6} = \Phi^{-1}(0.9522)$

$\frac{82 - \mu}{6} = 1.667$

$\therefore \mu = 71.998$

$= 72.0$ cm (3 s.f)