# Chapter ( 5 )

# Probability Distribution

* Probability : When an event may happen in $(x)$ ways out of a total of $(n)$ possible equally likely ways, the probability of occurrence (success) is given by :

$$p = Pr(E) = \frac{x}{n}$$

Hence the prob. Of non-occurrence (failure) is :

$$q = \Pr(\tilde{E}) = \frac{n-x}{n} = 1 - \frac{x}{n} = 1 - p$$

Thus $p + q = 1$

## Discrete Prob. Distribution :

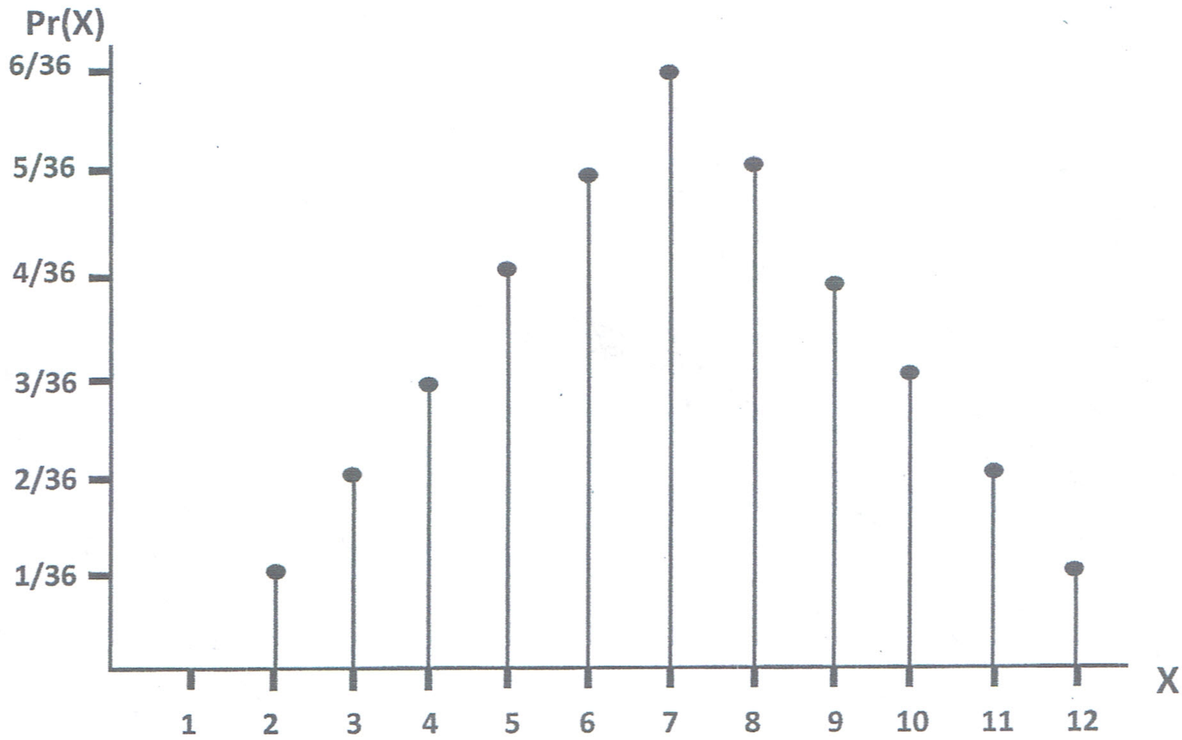
If a variable X may assume a set of discrete values $(x_i)$ with respective prob. $p_i$ , where $\sum p_i = 1$ , this defines a discrete prob. distribution for X.

### Example :

Let X be the sum of points obtained on a throw of two dice. The prob. or frequency distribution is given as :

| X : | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Pr(x): | $\frac{1}{36}$ | $\frac{2}{36}$ | $\frac{3}{36}$ | $\frac{4}{36}$ | $\frac{5}{36}$ | $\frac{6}{36}$ | $\frac{5}{36}$ | $\frac{4}{36}$ | $\frac{3}{36}$ | $\frac{2}{36}$ | $\frac{1}{36}$ |

$$\sum \Pr(X) = 1$$




* Relative frequency distribution of (N) throws is thus related to a sample of size (N) drawn out of an infinite population.

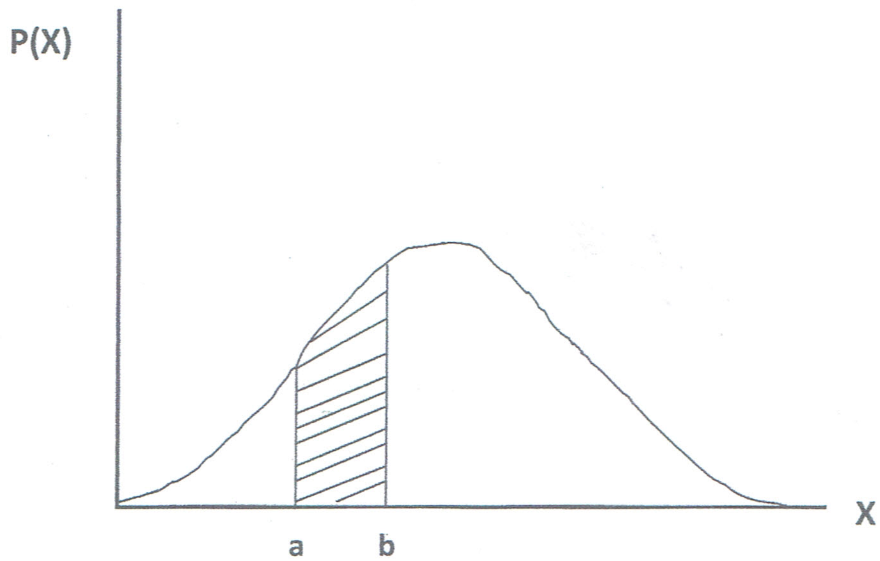
As $N \rightarrow \infty$ , the relative freq. dist. Approaches the prob. dist. Of the population .

## Continuous Probability Distribution :

* If a variable X may assume a continuous set of values, the prob. dist. Is a frequency curve where p(x)=fr .

Total area under the curve = $\sum fr = \sum p = 1$ .

* Prob. that X may lie between a and b ; Pr [ a < X < b ] = area under curve from a to b .




## The Normal Distribution

* The normal distribution is the most important of all probability distribution. It is applied directly to many practical problems, and several very useful distributions are based on it .

It is some times called the Gaussin dist. .

## Characteristics :

Many empirical freq. dist. Have the following characteristics :

1. They are approximately symmetrical, and the mode is close to the centre of the dist.
2. The mean, median, and mode are close together.
3. The shape of the dist. Can be approximated by a bell.

*** The prob. density function for the normal dist. Is given by:**

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

Where :

$\mu$ : is the mean of the theoretical dist.

$\sigma$ : is the standard deviation, and $\pi = 3.14$

*** This function extends from $-\infty$ to $\infty$**

Let $z = \dfrac{X-\mu}{\sigma}$ ,

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

* The total area bounded by the curve and the X axis is one. Hence, the area under the curve between X=a and X=b , Where a < b represent the prob. that X lies between a and b.

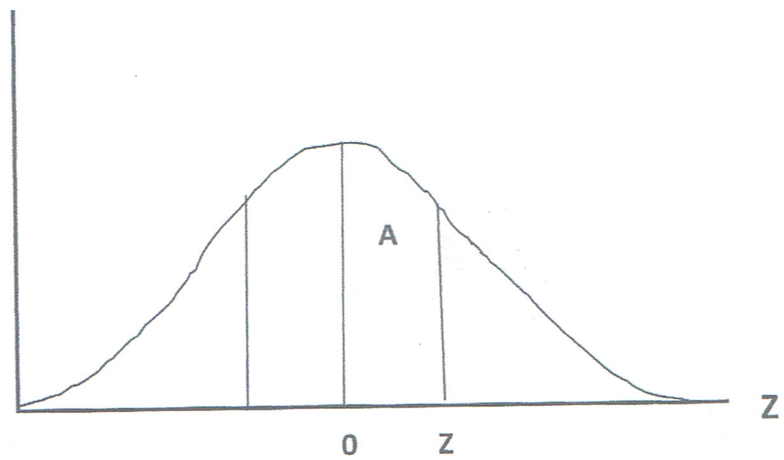
* Areas under the normal dist. Curve between 0 and Z are given in a :

Table (*). (Given below)

The prob. that Z lies between 0 and Z :

$Pr[0 < Z < z] = A$

From table (*) the area between any two ordinates can be found by using the symmetry of the curve about Z=0 .




*** Some properties of the normal dist. :**

$Mean = \mu$

$Variance = \sigma^2$

$Standard\ dev. = \sigma$

$$Mean\ dev. = \sigma\sqrt{\frac{2}{\pi}} = 0.7979\,\sigma$$

*** Areas under the normal curve :**

When $Pr[0 < Z < Z_1] = A$

$Pr[-Z_1 < Z < 0] = A$ (symmetrical curve).

$Pr[Z_1 < Z] = 0.5 - A$
$Pr[-Z_1 < Z] = 0.5 + A$
} Tatal area=1 so that area from $0 \rightarrow \infty$ is 0.5

When $Z_1$ and $Z_2$ are of same signs :

$Pr[Z_1 < Z < Z_2] = A_{Z_2} - A_{Z_1}$

When $Z_1$ and $Z_2$ are of different signs :

$$Pr[Z_1 < Z < Z_2] = A_{Z_2} + A_{Z_1}$$

* For bound of measurements, the bound in actual value is $\mp 0.5$ units in L.S.D.

**Example :**

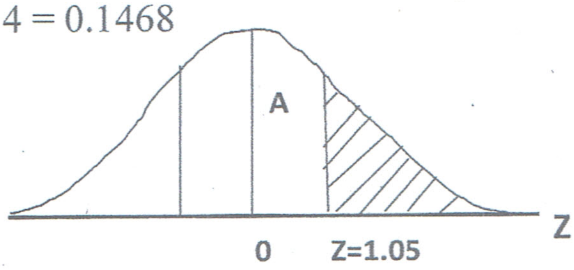
For measurements of $\mu = 160$ , $\sigma = 10$ , obtain the following :

1. Pr[X greater than 170] = Pr[X >170]

$$Z = \frac{X-\mu}{\sigma} = \frac{170.5-160}{10} = 1.05$$

$\therefore$ Pr[Z > 1.05]= = 0.5 − 0.35314 = 0.1468

Where $A_{Z=1.05} = 0.35314$

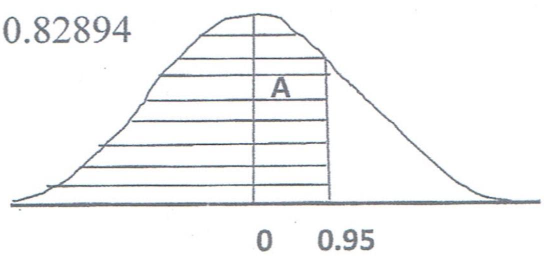



2. Pr[X less than 170] = Pr[X<170]

$x = 169.5 \ \rightarrow Z = 0.95$

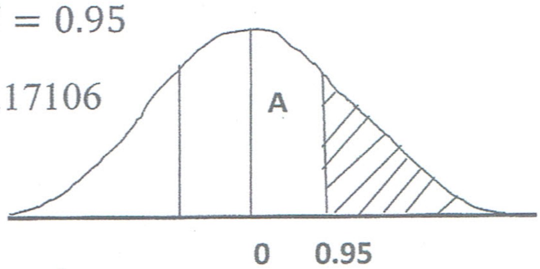
Pr [ Z < 0.95] = 0.5 + 0.32894 = 0.82894

Where $A_{Z=0.95} = 0.32894$



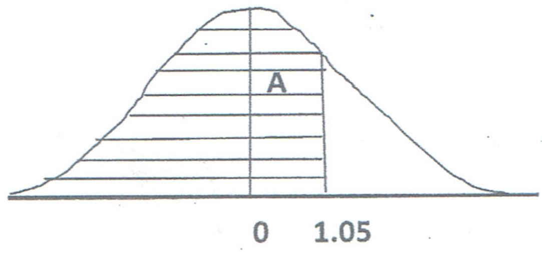

3. Pr[X ≥ 170] $\rightarrow x = 169.5 \ \rightarrow Z = 0.95$

Pr [ Z > 0.95] = 0.5 − 0.32894 = 0.17106



4. $\Pr[X \leq 170] \rightarrow x = 170.5 \quad \rightarrow Z = 1.05$

$\Pr[Z > 1.05] = 0.5 + 0.35314 = 0.85314$




## Linear inter polation :

When Z lies between successive $Z_1$ and $Z_2$ with respective $A_1$ and $A_2$ , A is obtained by linear interpolation "

$$\frac{Z - Z_1}{Z_2 - Z_1} = \frac{A - A_1}{A_2 - A_1} \qquad , Z_1 < Z < Z_2$$

e.g. for $\quad Z_1 = 2.32 \qquad A_1 = 0.48983$

$\qquad Z_2 = 2.33 \qquad A_2 = 0.49010$

Then when $Z = 2.327 \rightarrow A = ?$

$$A = \frac{Z - Z_1}{Z_2 - Z_1}(A_2 - A_1) + A_1 = 0.49002$$

**Example 1)**

For a measurement of size (N)=500

$\mu = 151$ , $\sigma = 15$ , assuming normal dist. Find how many measurements :

a ) between 120 and 155 = $\Pr[120 \leq X \leq 155]$

$x_1 = 119.5 \rightarrow z_1 = -2.1 \rightarrow A_1 = 0.4821$

$x_2 = 155.5 \rightarrow z_2 = 0.30 \rightarrow A_2 = 0.1179$

Pr $[-2.1 < Z < 0.3] = 0.4821 + 0.1179 =$ No. of meas. $= 500[0.4821+0.1179]=300$

b ) more than 185 = Pr[Z>185]

$x = 185.5 \rightarrow Z = 2.3 \rightarrow A = 0.4893$

Pr $[Z > 2.3] = 0.5 - 0.4893 =$ No. of meas. $= 500[0.51-0.4893]=5.$

c ) Less than 128 = Pr [X < 128]

$x = 127.5 \rightarrow Z = -1.57 \rightarrow A = 0.4418$

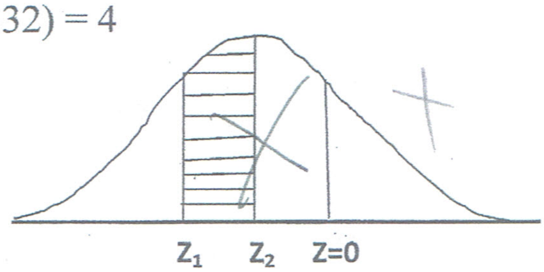
No. of meas. $= 500[0.5-0.4418]=29$

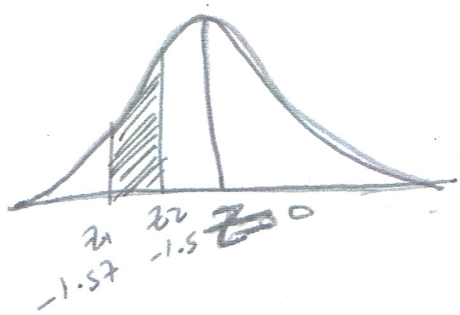
d ) equal to 128 = Pr [X=128]

$x_1 = 127.5 \rightarrow z_1 = -1.57 \rightarrow A_1 = 0.4418$

$x_2 = 128.5 \rightarrow z_2 = -1.5 \rightarrow A_2 = 0.4332$

Pr $[-1.57 < Z < -1.5] = 0.4418-0.4332 =$ No. of meas. $= 500(0.4418-0.4332) = 4$



e ) Less than or equal to 128 = Pr [X≤128]

$x = 128.5 \rightarrow Z = -1.5 \rightarrow A = 0.4332$

No. $= 500[0.5-0.4332]=33$

f ) Less than or equal to 185 = Pr [X≤185]

$x = 185.5 \rightarrow Z = 2.3 \rightarrow A = 0.4893$

No. $= 500[0.5 + 0.4893] = 495$

**Example 2 )**

For a sample of washers produced by a machine the mean inside dia. ($\mu$) is 5.02 mm and the standard deviation is 0.05 mm. The max. useful tolerance in the dia. Is 4.96 to 5.08 mm, otherwise the washers are considered defective. Determine % of defective washers.

Solu. )

Pr of max. to lerance = Pr $(4.96 \leq X \leq 5.08)$

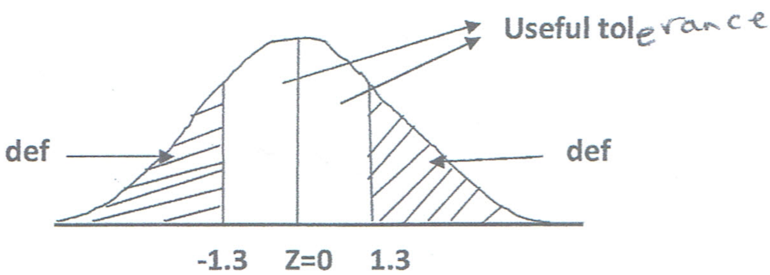
One unit in L.S.D. = 0.01

$x_1 = 4.955 \rightarrow z_1 = -1.3 \rightarrow A_1 = 0.4032$

$x_2 = 5.085 \rightarrow z_2 = +1.3 \rightarrow A_2 = 0.4032$

Pr $[-1.3 < Z < 1.3] = 2 * 0.4032 = 0.8064$

$\therefore$ % of defective washers = (1−0.8064) * = 19.4 %




**Example 3 )**

Out of a large No. of examination applicant a sample of size 50 gave a mean mark of 64 and a standard dev. of 14 . What is the expected % of applicants achieving a min. pass mark of 50 ?

Solu. :

one unit = 1

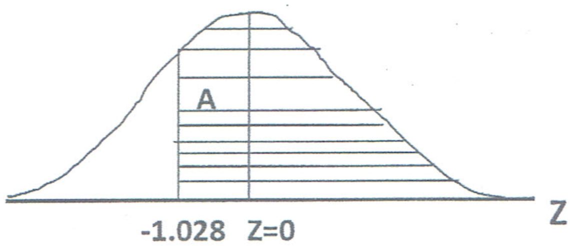
$\mu = 64$

$$\sigma = S\sqrt{\frac{N}{N-1}} = 14\sqrt{\frac{50}{50-1}} = 14.1$$

Pr [app. Have a min. pass mark of 50]= Pr [50≤X]

$x = 49.5 \rightarrow Z = -1.028 \rightarrow A = 0.3480$

Pr [$-1.028 < Z$] = 0.348 + 0.5 = 0.848 = 84.8 %




**Example 4 )**

The strength of individual bars made by a certain manufacturing process are approximate normally distributed with mean 28.4 and standard dev. 2.95 . To ensure safety, a customer requires at least 95% of the bars to be stronger than 24.0 . (one unit = 0.1)

a ) Do the bars meet the specification ?

b ) By improved manufacturing techniques, the manufacturer make the bars more uniform (that is, decrease the standard dev.) what value of standard dev. will just meet the specification if the mean stays the same ?

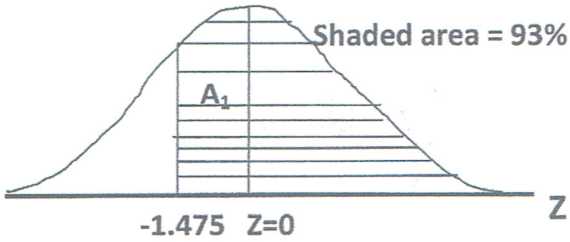
**Solu. :**

Pr [X > 24.0]→ $Z_1 = \frac{X-\mu}{\sigma} = \frac{24.05-28.4}{2.95}$

$$Z_1 = -1.475 \rightarrow A_1 = 0.4299$$

Pr [Z > - 1.475] = 0.5 + 0.4299 = 0.9299 ≃93%

Since (93%) less than 95% , the bars do not meet the specification.



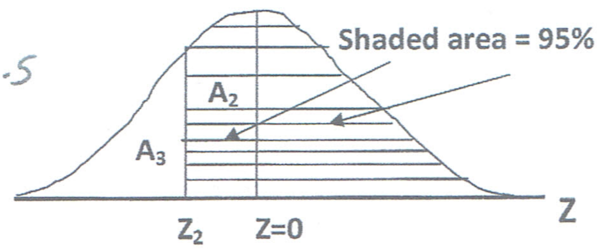

b ) The specification is at least 95% of bars > 24.0

$A_3 = 1 - 0.95 = 0.05$

$\therefore \ A_2 = 0.5 - 0.05 = 0.45$

oR: $A_2 = 0.95 - 0.5$

$A_2 = 0.45$




At $A_2 = 0.45 \rightarrow Z_2 = -1.645$ (from table)

$$Z_2 = \frac{X - \mu}{\sigma} \rightarrow -1.645 = \frac{24.05 - 28.4}{\sigma}$$

$\sigma = 2.644$ (if the $\sigma$ can be reduced to 2.644 while keeping the mean constant, the specification will just be met)