Continuous Probability distribution:

Normal distribution: (also known as Gaussian distribution)

The most important cont. prob. distribution in the field of statistics is normal distribution and its graph is called the normal curve.

A continuous r.v X having the bell shaped distribution is called normal r.v.

Definition of Normal distribution: Thus the density function of the normal r.v X with mean μ and variance σ^2 is given by

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

where π=3.14159 and e=2.71828

Once μ and σ are specified, the normal curve is completely determined.

Some characteristics of normal distribution:

1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurred at $x=\mu$.

2.The curve is symmetric about a vertical axis through the mean μ.

3. The curve has its points of inflection at $x=\mu \pm \sigma$; it is concave downward if $\mu - \sigma < x < \mu + \sigma$ and is concave upward otherwise.

4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.

5. The total area under the curve and above the horizontal axis is equal to 1.

Now it can be shown that the parameters μ and σ^2 are indeed the mean and variance of the normal distribution.

To evaluate the mean of Normal distribution:

We write $E(X)=\int_{-\infty}^{\infty} x\, n(x;\mu,\sigma)\,dx$

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

$$=\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} x \; e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \qquad \text{Let } \frac{x-\mu}{\sigma}=z$$

$$=\frac{1}{\sqrt{2\pi}\sigma} \lim_{\substack{A\to-\infty \\ B\to\infty}} \int_A^B (\mu+\sigma z)\, e^{-\frac{z^2}{2}} \sigma dz$$

$$=\lim_{\substack{A\to\infty\\B\to\infty}} [\frac{\mu}{\sqrt{2\pi}}\int_A^B e^{-\frac{z^2}{2}}dz + \frac{\sigma}{\sqrt{2\pi}}\int_A^B ze^{-\frac{z^2}{2}}dz \ ]$$

In the first integration we see that $\mu.\frac{1}{\sqrt{2\pi}.1}\int_{-\infty}^{\infty} e^{-\frac{(z-0)^2}{2.1^2}}$ dz which is the μ times the area under the normal curve with mean 0 and variance 1 and hence $\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{z^2}{2}}dz$ =1

Therefore the first integral reduces to μ.

Next for the 2nd integral $\int z\, e^{-\frac{z^2}{2}}\, dz$

$$=-e^{-\frac{z^2}{2}}$$

Therefore $-[e^{-\frac{z^2}{2}}\ ]_{-\infty}^{\infty}$=-0

The value of the 2nd integral is zero

Therefore E(X) =μ which is the mean of the normal distribution.

To evaluate the variance of normal distribution:

Var. = E[$(X-\mu)^2$]=$\int_{-\infty}^{\infty}(x-\mu)^2\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$dx.

Putting $\frac{x-\mu}{\sigma}$=z i,e dx=σdz

$$=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\sigma^2 z^2\, e^{-\frac{z^2}{2}}\, dz = \frac{1}{\sqrt{2\pi}}\lim_{\substack{A\to-\infty\\B\to\infty}}\int_A^B \sigma^2 z^2\, e^{-\frac{z^2}{2}}dz$$

$$=\lim_{\substack{A\to-\infty\\B\to\infty}} [\frac{\sigma^2}{\sqrt{2\pi}}\int_A^B z^2 e^{-\frac{z^2}{2}}dz \ ]= \frac{\sigma^2}{\sqrt{2\pi}}\lim_{\substack{A\to-\infty\\B\to\infty}} [\int_A^B z.ze^{-\frac{z^2}{2}}dz\ ]$$

$$=\frac{-\sigma^2}{\sqrt{2\pi}}\lim_{\substack{A\to-\infty\\B\to\infty}}[z.e^{-\frac{z^2}{2}}\ ]_A^B + \frac{\sigma^2}{\sqrt{2\pi}}\lim_{\substack{A\to-\infty\\B\to\infty}}\int_A^B e^{-\frac{z^2}{2}}\ \text{dz}$$ (Using integration by parts)

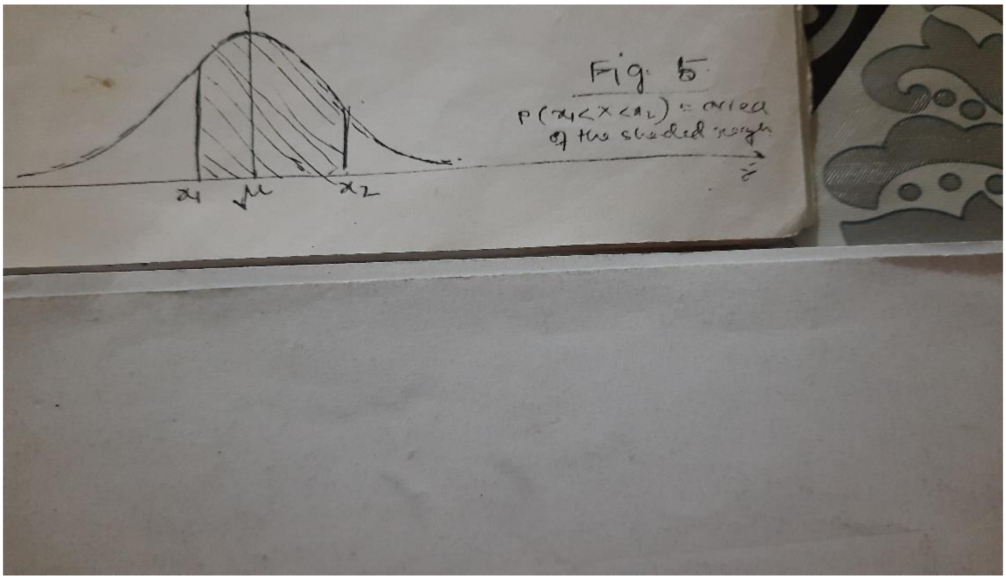
=0 (using L' Hospital's rule) + σ^2.1 =σ^2

The variance of the normal distribution is σ^2.

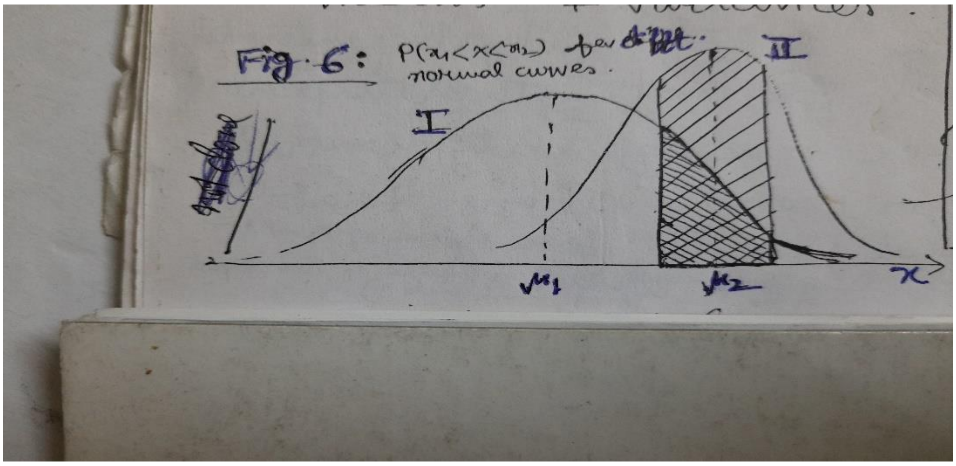
Area under the Normal Curve:

The curve of any continuous prob. distribution or density function is constructed so that the area under the curve bdd. by the two ordinates $x=x1$ and $x=x2$ equals the prob. that the r.v X assumes a value between $x=x1$ and $x=x2$. Thus for normal curve $P(x1<X<x2)=\int_{x1}^{x2} n(x;\mu,\sigma)dx$

$=\frac{1}{\sqrt{2\pi}\sigma}\int_{x1}^{x2} e^{-(x-\mu)^2/2\sigma^2}dx$ is represented by the area of the shaded region (Fig.-5).



The areas also depends on the values μ and σ. This is found in Fig-6 where we have shaded the region corresponding to $P(x1<X<x2)$ for two curves with different means and variances.



Here the P(x1<X<x2) is indicated by the darker shaded area where X is the r.v describing the distribution I. If X is the r.v describing distribution II then P(x1<X<x2) is given by the entire shaded region.

Obviously the two shaded region are different in size and therefore prob. associated with each distribution will be different for two given values.

It is to be noted that all the observations of any normal r.v. X can be transformed to a new set of observations of a normal r.v. Z with mean zero and variance 1. This can be done by means of the transformation

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

Therefore

P(x1<X<x2)=$\int_{x1}^{x2} n(x;\mu,\sigma)$dx= $\frac{1}{\sqrt{2\pi}\sigma}\int_{x1}^{x2} e^{-(x-\mu)^2/2\sigma^2}dx$ =$\frac{1}{\sqrt{2\pi}\sigma}\int_{z1}^{z2} e^{-(z)^2/2}\sigma dz$ =$\frac{1}{\sqrt{2\pi}}\int_{z1}^{z2} e^{-(z)^2/2}dz$, where z=$\frac{x-\mu}{\sigma}$.

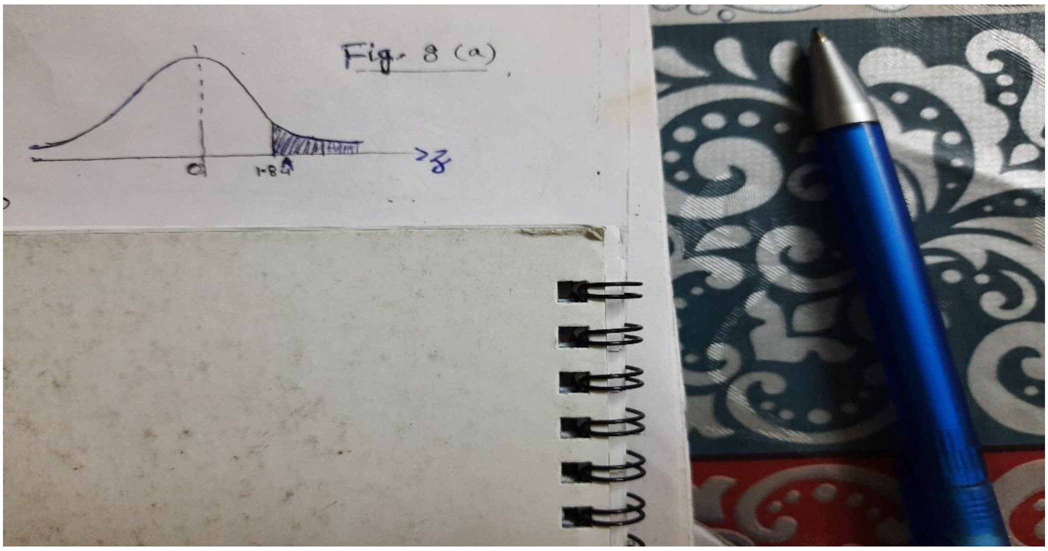
=$\frac{1}{\sqrt{2\pi}.1}\int_{z1}^{z2} e^{-(z-0)^2/2.1^2}dz$=$\int_{z1}^{z2} n(z;0,1)$dz=P(z1<Z<z2).

where z is a normal r.v with mean 0 and variance 1.

Thus the distribution of a normal r.v with mean zero and variance 1 is called a standard normal distribution & Z →standard normal variate.

Problem 1.: Given a standard normal distribution, find the area under the curve that lies (a) to the right of z=1.84 and (b) between z= -1.97 and z=0.86.

Sol: (a) Area corresponding to z=1.84 up to the point A area to the right of z=1.84 is equal to 1 minus the area in the given table to the left of z=1.84 namely 1-0.9671=0.0329 (Ans)



(b) Area corresponding to z=-1.97 is equal to .0244 which is upto the point A . Also area corresponding to z=0.86 is equal to 0.8051 which is upto the point B. Therefore area between z=-1.97 and z=0.86 is given by the area to the left of z=0.86 minus the area to the left of z=-1.97 and is equal to 0.8051-0.0244=0.7807 (Ans)