NORMAL DISTRIBUTION

Let μ and σ be two arbitrary constants such that $-\infty < \mu < \infty$ and $\sigma > 0$, then the probability distribution for which

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

is the density function is called the normal distribution and the corresponding random variable X is called the normal variate.

Clearly,

i) For any X, $N(\mu, \sigma) \geq 0$

ii) $\int_{-\infty}^{\infty} N(\mu, \sigma)\, dx = 1$

Verification

$$\mathrm{I} = \int_{-\infty}^{\infty} N(\mu, \sigma)\, dx$$

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

Let $\frac{x-\mu}{\sigma\sqrt{2}} = \theta$

$$dx = \sigma\sqrt{2}\, d\theta$$

θ varies from $-\infty$ *to* ∞

$$\mathrm{I} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\theta^2}\, \sigma\sqrt{2}\, d\theta$$

$$\mathrm{I} = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\theta^2}\, d\theta$$

$\mathrm{I} = \frac{1}{\sqrt{\pi}}\, 2 \int_0^{\infty} e^{-\theta^2}\, d\theta$ $\left(\because e^{-\theta^2}\right.$ is even function)

$$\mathrm{I} = \frac{2}{\sqrt{\pi}} \left(\frac{\sqrt{\pi}}{2}\right) = 1$$

$\left(\because \int_0^{\infty} e^{-\theta^2}\, d\theta = \frac{\sqrt{\pi}}{2}\right.$ **by gamma functions**$\Big)$

$$\therefore \boxed{\int_{-\infty}^{\infty} N(\mu, \sigma)\, dx = 1}$$

Note:

Normal distribution is limiting case of Binomial distribution when n is very large and neither p nor q is very small.

Mean and Variance of Normal distribution

$$\text{Mean} = \int_{-\infty}^{\infty} x.N(\mu,\sigma)\,dx$$

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

Let $\frac{x-\mu}{\sigma\sqrt{2}} = \theta$ so that $x = \mu + \sigma\sqrt{2}\theta$

$$dx = \sigma\sqrt{2}\,d\theta$$

θ varies from $-\infty$ to ∞

$$\textbf{Mean} = \frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}(\mu+\sigma\sqrt{2}\theta)\,e^{-\theta^2}\,\sigma\sqrt{2}\,d\theta$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}(\mu+\sigma\sqrt{2}\theta)\,e^{-\theta^2}\,d\theta$$

$$= \frac{\mu}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\theta^2} d\theta + \frac{\sigma\sqrt{2}}{\sqrt{\pi}}\int_{-\infty}^{\infty} \theta e^{-\theta^2}\ d\theta$$

The value of second integral is zero because we know that $\int_{-a}^{a} f(x)dx = 0$ if $f(x)$ is odd function of x. In this case $e^{-\theta^2}$ is even function but θ is odd function and hence $\theta e^{-\theta^2}$ is odd function.

$$\therefore \textbf{Mean} = \frac{\mu}{\sqrt{\pi}}\int_{-\infty}^{\infty} e^{-\theta^2} d\theta$$

$$= \frac{\mu}{\sqrt{\pi}}\left(2\int_{0}^{\infty} e^{-\theta^2}\ d\theta\right)$$

$$= \frac{\mu}{\sqrt{\pi}}\left(2\frac{\sqrt{\pi}}{2}\right)\ \left(\because \int_{0}^{\infty} e^{-\theta^2}\ d\theta = \frac{\sqrt{\pi}}{2}\right)$$

$$\therefore \boxed{\textbf{Mean} = \mu}$$

Conclusion:

The parameter μ present in the function $N(\mu, \sigma)$ is mean of the Normal distribution.

Variance

Let V be variance of the Normal distribution.

Then,

$$\mathbf{V} = \int_{-\infty}^{\infty} (x-\mu)^2 N(\mu,\sigma)\, dx$$

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

Let $\frac{x-\mu}{\sigma\sqrt{2}} = \theta$

$$dx = \sigma\sqrt{2}\, d\theta$$

θ varies from $-\infty$ to ∞

$$\mathbf{V} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\sigma\sqrt{2}\,\theta\right)^2 e^{-\theta^2}\, \sigma\sqrt{2}\, d\theta$$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\,\theta)^2 e^{-\theta^2}\; d\theta$$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}}\left(2\int_0^\infty (\theta)^2 e^{-\theta^2}\ d\theta\right)$$

$$\left(\because \theta^2 e^{-\theta^2} \text{is even function}\right)$$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}}\left(\int_0^\infty \theta e^{-\theta^2}(2\theta)\,d\theta\right)$$

Let $\theta^2 = t$ then $2\theta d\theta = dt$

t also varies from $0\ to\ \infty$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}}\left(\int_0^\infty t^{\frac{1}{2}} e^{-t}\,dt\right)$$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}+1\right) \quad \left(\because \Gamma(n+1) = \int_0^\infty t^n e^{-t}\,dt\right)$$

$$\mathbf{V} = \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{2\sigma^2}{\sqrt{\pi}}\frac{1}{2}\sqrt{\pi} = \sigma^2$$

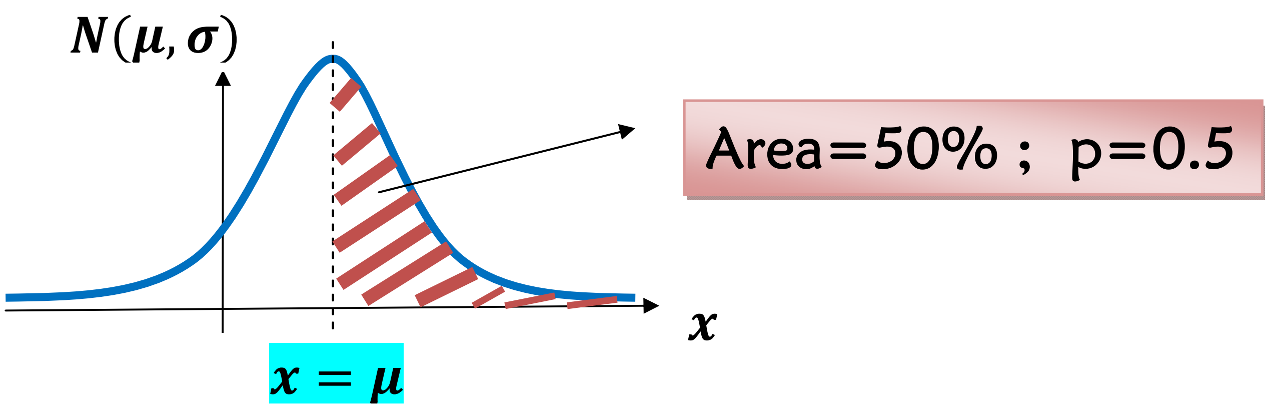
Thus, Varience $= \sigma^2$ & S.D $= \sigma$

Conclusion:

The parameter $\boldsymbol{\sigma}$ present in the function $\boldsymbol{N(\mu,\sigma)}$ is standard deviation of the Normal distribution.

Normal curve

The graph of the function $\boldsymbol{N(\mu,\sigma)}$ known as NORMAL CURVE. It is bell shaped as shown below which is symmetric about the line $\boldsymbol{x=\mu}$.



The line $\boldsymbol{x=\mu}$ divides the total area under the curve (Area=1) into two equal parts.

Note:

If X is continuous random variable, then $\boxed{P(a \leq X \leq b) = \int_a^b f(x)\,dx}$ represents the area under the curve $y = f(x)$ between $x = a$ and $x = b$.

However, if X is a normal variate, then we can find such areas without calculus.

Standard Normal Distribution

The Normal distribution for which $\mu = 0$ and $\sigma = 1$ is known as Standard Normal distribution.

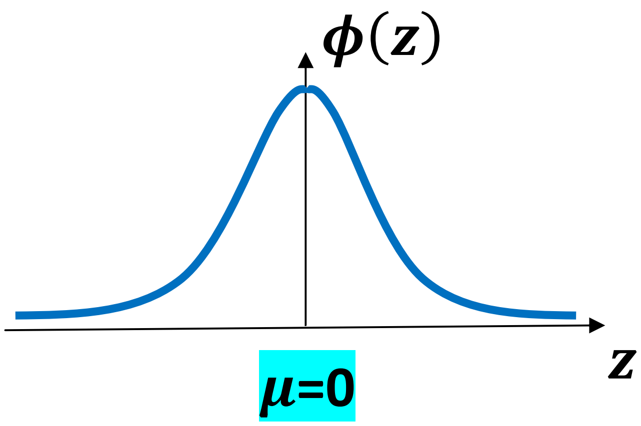
Suppose X is normal variate, then the standard normal variable Z corresponding to X is defined as

$$\boxed{Z = \frac{X - \mu}{\sigma}}$$

The density function for the standard normal distribution is

$$\boxed{\phi(z) = N(0,1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}}$$

Standard Normal curve



Evaluating Normal Probabilities

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

If X is normal variable, then

$$P(a \leq X \leq b) = \int_a^b N(\mu, \sigma)\, dx$$

$$P(a \leq X \leq b) = \int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}\, dx$$

Let $Z = \frac{X-\mu}{\sigma}$ be the standard normal variable.

Let $Z = z_1$ when $X = a$

$Z = z_2$ when $X = b$

Then,

$P(a \leq X \leq b) = P(z_1 \leq Z \leq z_2)$

$$= \int_{z_1}^{z_2} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz$$

$$\therefore \boxed{P(z_1 \leq Z \leq z_2) = \int_{z_1}^{z_2} \phi(z)\, dz}$$

Note:

1. $\int_{z_1}^{z_2} \phi(z)\, dz$ represents area under standard normal curve between z_1 and z_2.

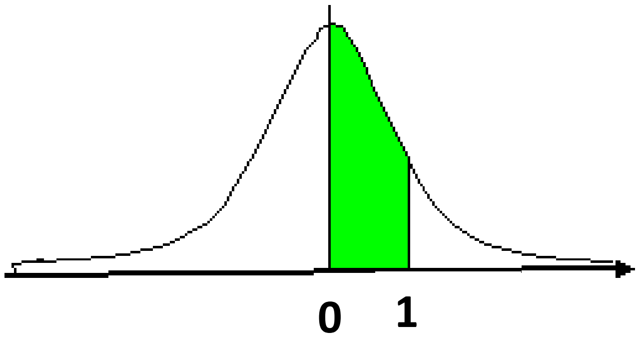
2. The integral $\boxed{A(z) = \int_0^z \phi(z)\, dz}$ represents the area under the standard normal from 0 to z.

3. $\boxed{P(0 \leq Z \leq c) = A(c)}$

Here, $A(c)$ can be obtained from standard normal table.

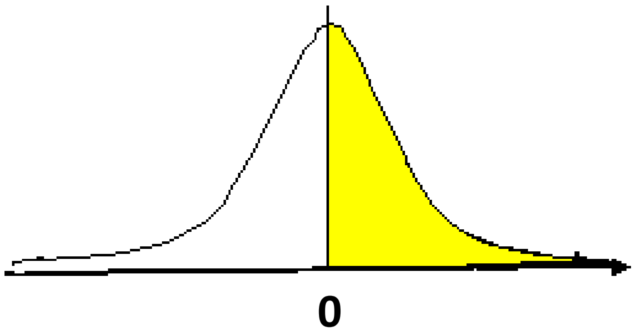
Example:

1



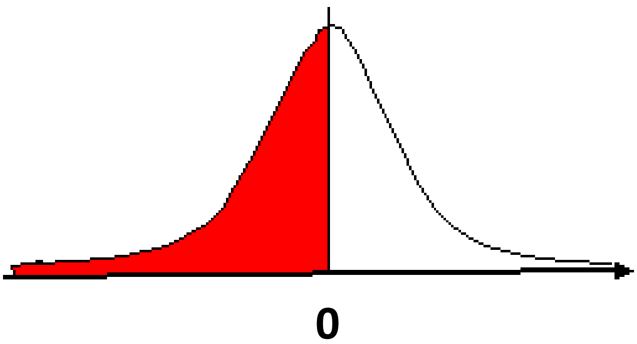
$P(0 \leq Z \leq 1) = A(1)$
$= 0.3413$

2



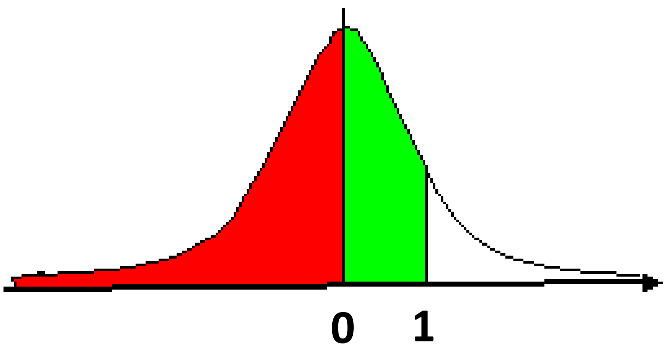
$P(0 \leq Z < \infty)$
$= 0.5$

3



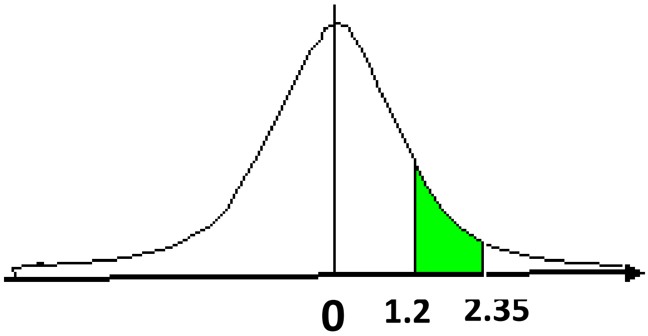
$P(-\infty < Z \leq 0)$
$= 0.5$

4



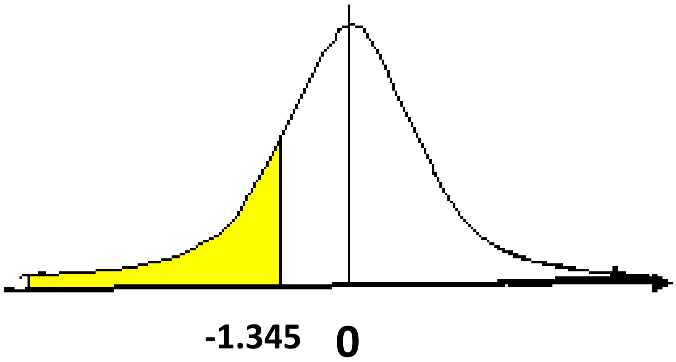
$P(-\infty < Z \leq 1)$
$= 0.5 + A(1)$
$= 0.5 + 0.3413$
$= 0.8413$

5



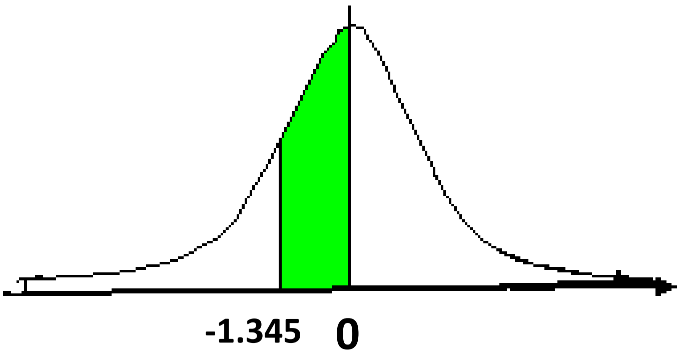
$$P(1.2 \le Z \le 2.35)$$
$$= A(2.35) - A(1.2)$$
$$= 0.4906 - 0.3849$$
$$=0.1057$$

6



$$P(-\infty < Z \le -1.345)$$
$$= P(1.345 < Z < \infty)$$
$$= 0.5 - A(1.345)$$
$$= 0.5 - 0.4099$$
$$= 0.0901$$

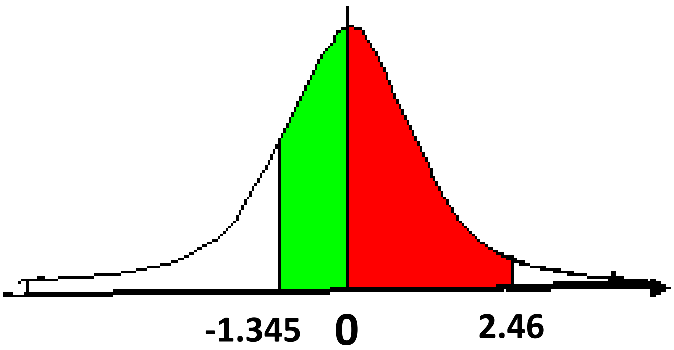
7



$$P(-1.345 \le Z \le 0)$$
$$= P(0 \le Z \le 1.345)$$

$$= A(1.345)$$
$$= 0.4099$$

8



$$P(-1.345 \le Z \le 2.46)$$

Standard Normal Distribution table

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Problems

1. Define pdf of normal distribution. For the standard normal distribution of a random variable z, evaluate the following:
   (i) $P\ (\ \mathbf{0 \leq z\ \leq 1.45}\ )$
   (ii) $P\ (\ \mathbf{-2.60 \leq z \leq 0})$
   (iii) $P\ (\ \mathbf{-3 \leq z\ \leq 2.65})$
   (iv) $P\ (\ \mathbf{1.25 \leq z\ \leq 2.1})$
   (v) $P\ (\ \mathbf{z \geq 1.7}\ )$
   (vi) $P(|\mathbf{z}| \leq \mathbf{1.85})$

Soln:

(i) $\boldsymbol{P}(\mathbf{0 \leq z\ \leq 1.45}) = \boldsymbol{A}(\mathbf{1.45}) =$

(ii) $P\ (\ \mathbf{-2.60 \leq z \leq 0})$
This is same as
$P(\ \mathbf{0 \leq z \leq 2.60})$ =A(2.6)=
(due to symmetry)

(iii) $P\ (\ \mathbf{-3 \leq z\ \leq 2.65})$=A(3)+A(2.65)
(iv) $P\ (\ \mathbf{1.25 \leq z\ \leq 2.1})$=A(2.1)-A(1.25)
(v) $P\ (\ \mathbf{z \geq 1.7}\ )$= 0.5-A(1.7)
(vii) $P(|\mathbf{z}| \leq \mathbf{1.85})$=P(-1.85<z<1.85)
=2A(1.85)

2. For the normal distribution with mean μ and standard deviation σ, evaluate the following.
   (i) $P(|X-\mu| \leq \sigma)$
   (ii) $P(|X-\mu| \leq 2\sigma)$
   (iii) $P(|X-\mu| \leq 3\sigma)$

Soln:

Wkt, if X is normal variate, then standard normal variable is $Z = \frac{X-\mu}{\sigma}$

(i) $P(|X-\mu| \leq \sigma) = P\left(\frac{|X-\mu|}{\sigma} \leq 1\right)$

$$= P\left(\left|\frac{X-\mu}{\sigma}\right| \leq 1\right)$$

$$= P(|Z| \leq 1)$$

$$= P(-1 \leq Z \leq 1)$$

$$= P(-1 \leq Z \leq 0) + P(0 \leq Z \leq 1)$$

Due to symmetry

$$P(-1 \leq Z \leq 0) = P(0 \leq Z \leq 1) = A(1)$$

$$\therefore P(|X-\mu| \leq \sigma) = 2A(1) = 2 \times 0.3413$$

$$\therefore \boxed{P(|X-\mu| \leq \sigma) = 0.6826}$$

(ii) $P(|X-\mu| \leq 2\sigma) = P\left(\frac{|X-\mu|}{\sigma} \leq 2\right)$

$$= P\left(\left|\frac{X-\mu}{\sigma}\right| \leq 2\right)$$

$$= P(|Z| \leq 2)$$

$$= P(-2 \leq Z \leq 2)$$

$$= 2P(0 \leq Z \leq 2)$$

$$= 2A(2) = 2 \times 0.4772$$

$$\therefore \boxed{P(|X-\mu| \leq 2\sigma) = 0.9544}$$

(iii) $P(|X-\mu| \leq 3\sigma) = P\left(\left|\frac{X-\mu}{\sigma}\right| \leq 3\right)$

$$= P(|Z| \leq 3)$$

$$= P(-3 \leq Z \leq 3)$$

$$= 2P(0 \leq Z \leq 3)$$

$$= 2A(3) = 2 \times 0.4987$$

$$\therefore \boxed{P(|X-\mu| \leq 3\sigma) = 0.9974}$$

3. For the normal distribution with mean 3 and standard deviation 5, evaluate the following probabilities:
(i)$P(x \geq 6)$ (ii) $P(|x| < 3)$ (iii)$P(|x| > 4)$

Soln:

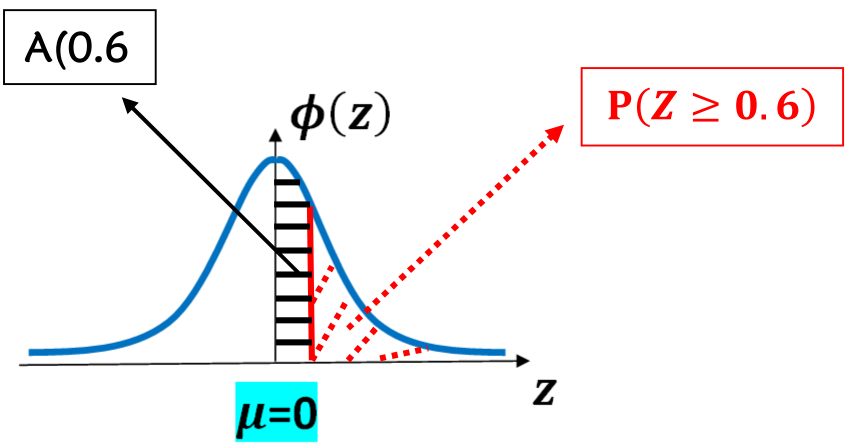
Given $\boldsymbol{\mu = 3 \;\&\; \sigma = 5}$

Wkt, if $\mathbf{X}$ is normal variate, then standard normal variable is $\boldsymbol{Z = \frac{X-\mu}{\sigma} = \frac{X-3}{5}}$

(i) $\mathbf{P(x \geq 6) = P\left(\frac{X-\mu}{\sigma} \geq \frac{6-\mu}{\sigma}\right)}$

$$\mathbf{= P\left(Z \geq \frac{6-3}{5}\right)}$$

$$\mathbf{= P(Z \geq 0.6) = 0.5 - A(0.6)}$$
$$\mathbf{= 0.5 - 0.2257 = 0.2743}$$



(ii) $P(|x| < 3) = \mathbf{P(-3 < X < 3)}$

$$= \mathbf{P}\left(\frac{-3-\mu}{\sigma} < \frac{X-\mu}{\sigma} < \frac{3-\mu}{\sigma}\right)$$

$$= \mathbf{P}\left(\frac{-3-3}{5} < Z < \frac{3-3}{5}\right)$$

$$= \mathbf{P(-1.2 < Z < 0)}$$

$$= \mathbf{P(0 < Z < 1.2)} \quad \textbf{d uetto symmetry}$$

$$= \mathbf{A(1.2)} = \mathbf{0.3849}$$

(iii) $\mathbf{P}(|x| > 4) =$