Computer Science & DA

Probability and Statistics

Continuous Random variable

Lecture No. 04

By- Dr. Puneet Sharma Sir

Recap of previous lecture

Topic Normal Distribution 1

Topics to be Covered

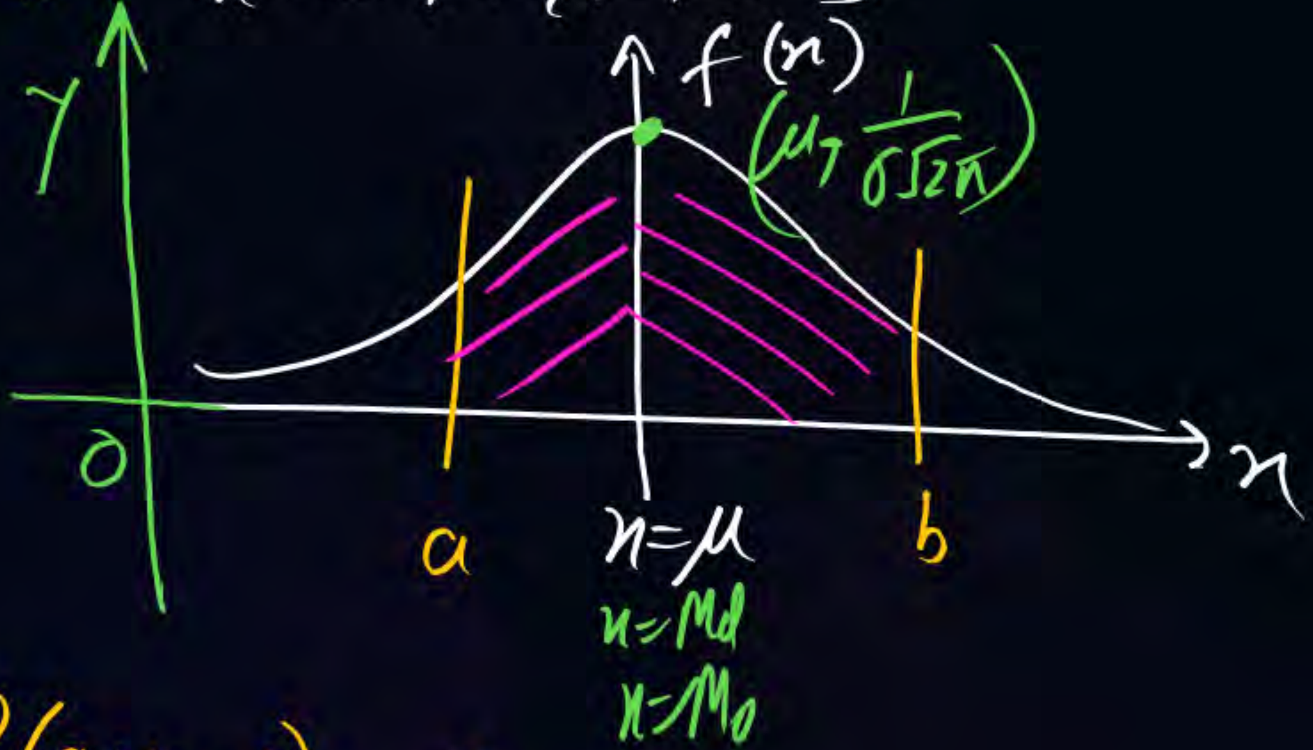
Topic Normal Distribution 2 -CLT, NST, CDF.

RECAP:-

N.R.V(x) eg: Height, wt, M. Dist, Intelligence

Def^n: $f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}$ = NENO

ie $x \sim N\{\mu, \sigma^2\}$



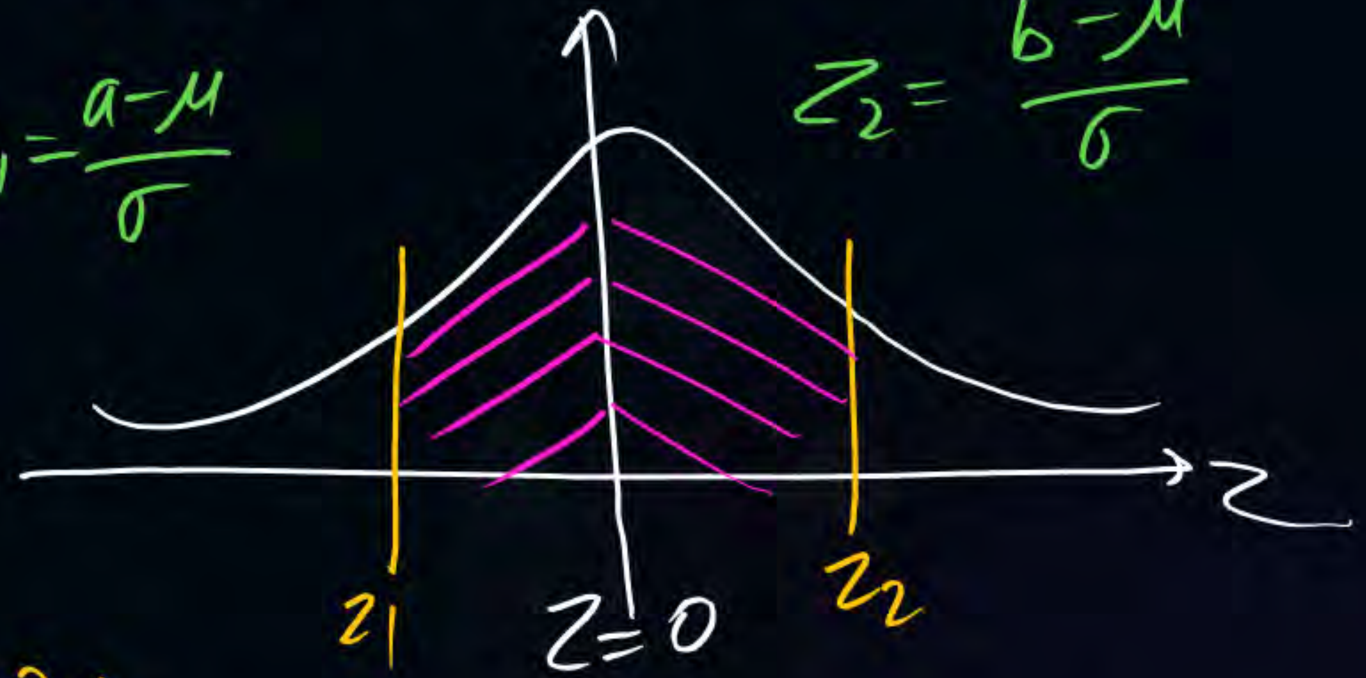
$P(a<x<b) = ?$

S.N.V(Z) → $\mu = 0, \sigma = 1$ & $Z = \frac{x-\mu}{\sigma}$

$f(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\}$ = Even func^n

ie $Z \sim N\{0, 1\}$

$z_1 = \frac{a-\mu}{\sigma}$ $\quad$ $z_2 = \frac{b-\mu}{\sigma}$



$P(a<x<b) = ? = P(z_1<z<z_2)$

= use N-Table

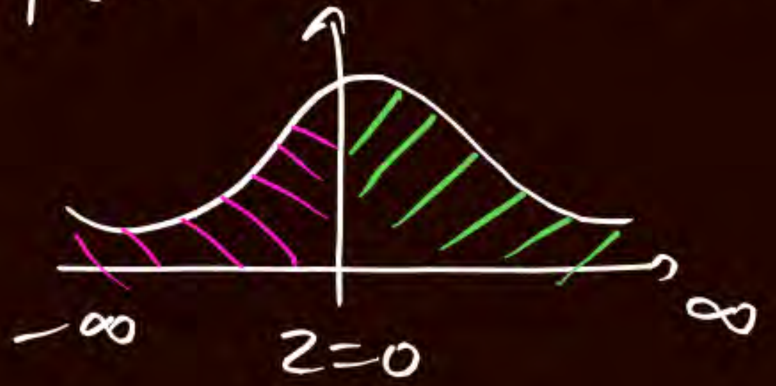
Symmetry of N. Curve :-

In General, N. Table starts from $z = 0$ and is defined for +ve values of z.

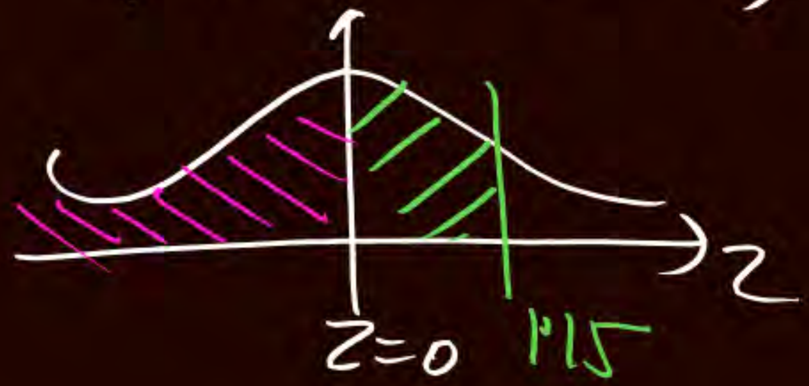
① $P(-\infty < z < \infty) = 1$ i.e Total area under N.C.

② $P(-\infty < z < 0) = 0.5$ i.e left Half area " "

③ $P(0 < z < \infty) = 0.5$ i.e Right " " "



④ $P(-\infty < z \le 1.15) = ?$ $= 0.5 + P(0 < z < 1.15)$



$= \frac{1}{2} + (\text{use N. Table})$

$= \frac{1}{2} + 0.3746$

⑤ $P(-1.6 < z < 1.6) = ?$ $= 2P(0 < z < 1.6)$

= use N. Table

$= 2(0.4452)$

⑥ $P(-3 < z \le 2) = ?$

$= P(-3 < z < 0) + P(0 < z \le 2)$

$= P(0 < z < 3) + P(0 < z \le 2)$

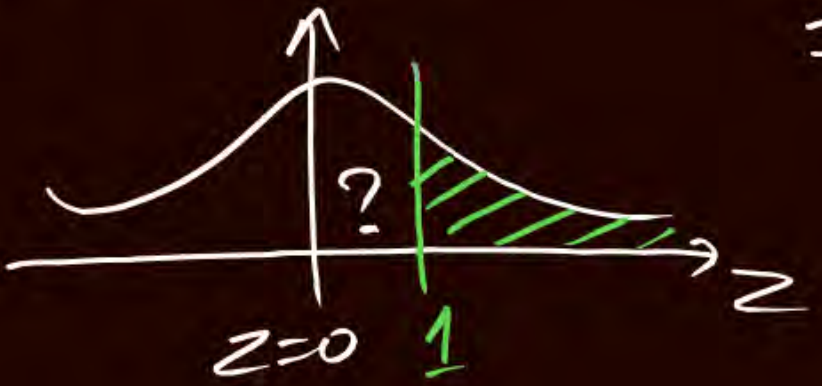
$= \left(\frac{0.997}{2}\right) + \left(\frac{0.955}{2}\right)$

$= ?$

⑦ $P(z \geq 1) = ?$ $= \text{Right Half area} - P(0 \leq z \leq 1)$

$= 0.5 - (0.3413)$

$= ?$



flowchart of Solving Questions →

Step 1 → First find x = {which is Required}

Step 2 → Convert x into z using $\frac{x-\mu}{\sigma} = z$

Step 3 → use concept of Symmetry

Step 4 → use N. Table.

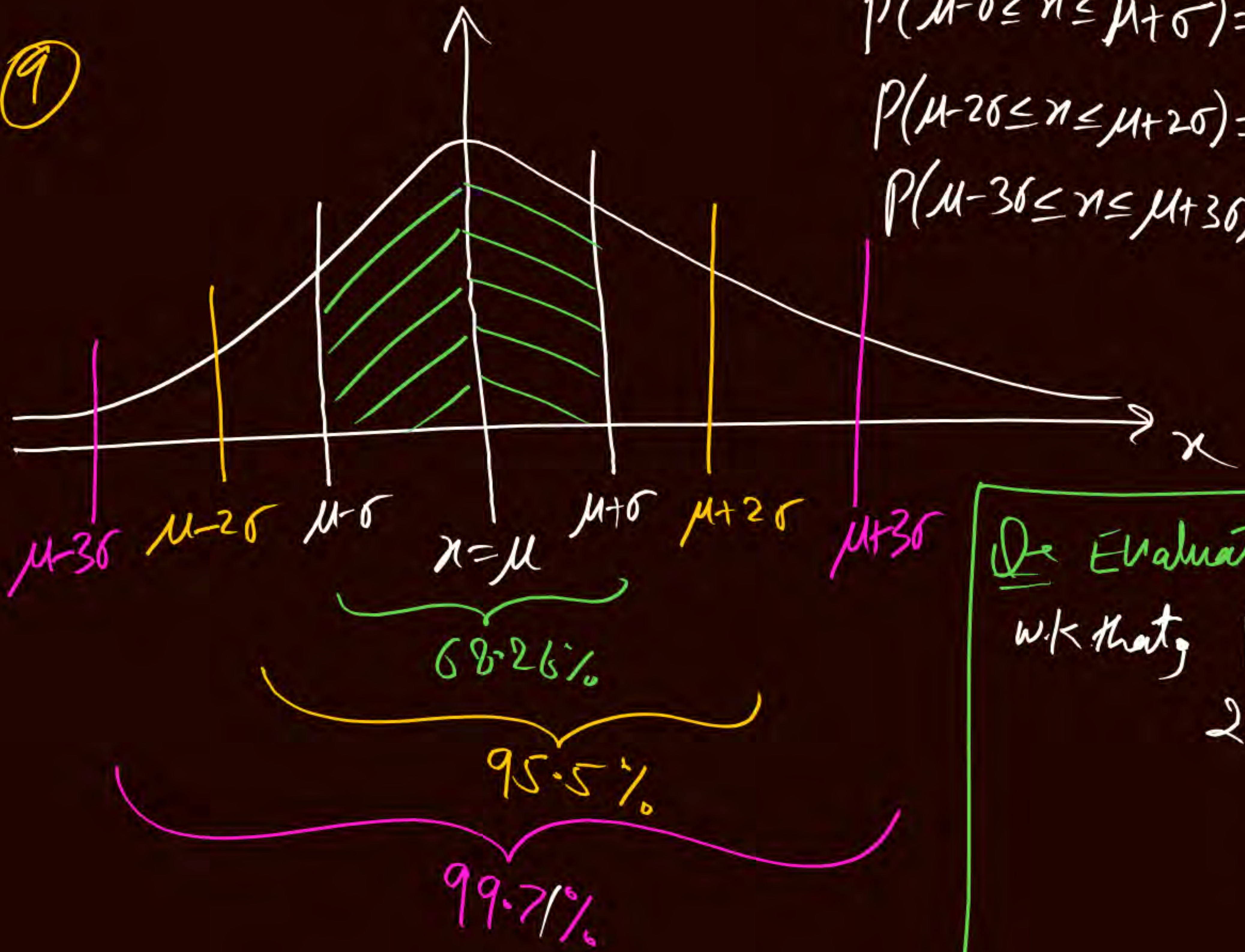
⑧ Skewness in N-Curve →

(No skewness) (Me=Md=Mo)

(-vely skewed) (Me<Md<Mo)

(+vely skewed) (Mo<Md<Me)

⑨



$P(\mu-\sigma \le x \le \mu+\sigma) = 0.6826 \quad \longrightarrow \quad P(-1 \le z \le 1)$

$P(\mu-2\sigma \le x \le \mu+2\sigma) = 0.955 \quad \xrightarrow[\sigma=1]{\mu=0} \quad P(-2 \le z \le 2)$

$P(\mu-3\sigma \le x \le \mu+3\sigma) = 0.9971 \quad \longrightarrow \quad P(-3 \le z \le 3)$

Q: Evaluate $P(0 < z < 1) = ?$

w.k. that, $P(-1 \le z \le 1) = 0.6826$

$2P(0 \le z \le 1) = 0.6826$

$P(0 \le z \le 1) = \frac{0.6826}{2} = 0.3413$

Qs $x \sim N\{102, 27^2\}$ then evaluate $P(90 \le x \le 102) = ?$

(a) 68%

(b) 50%

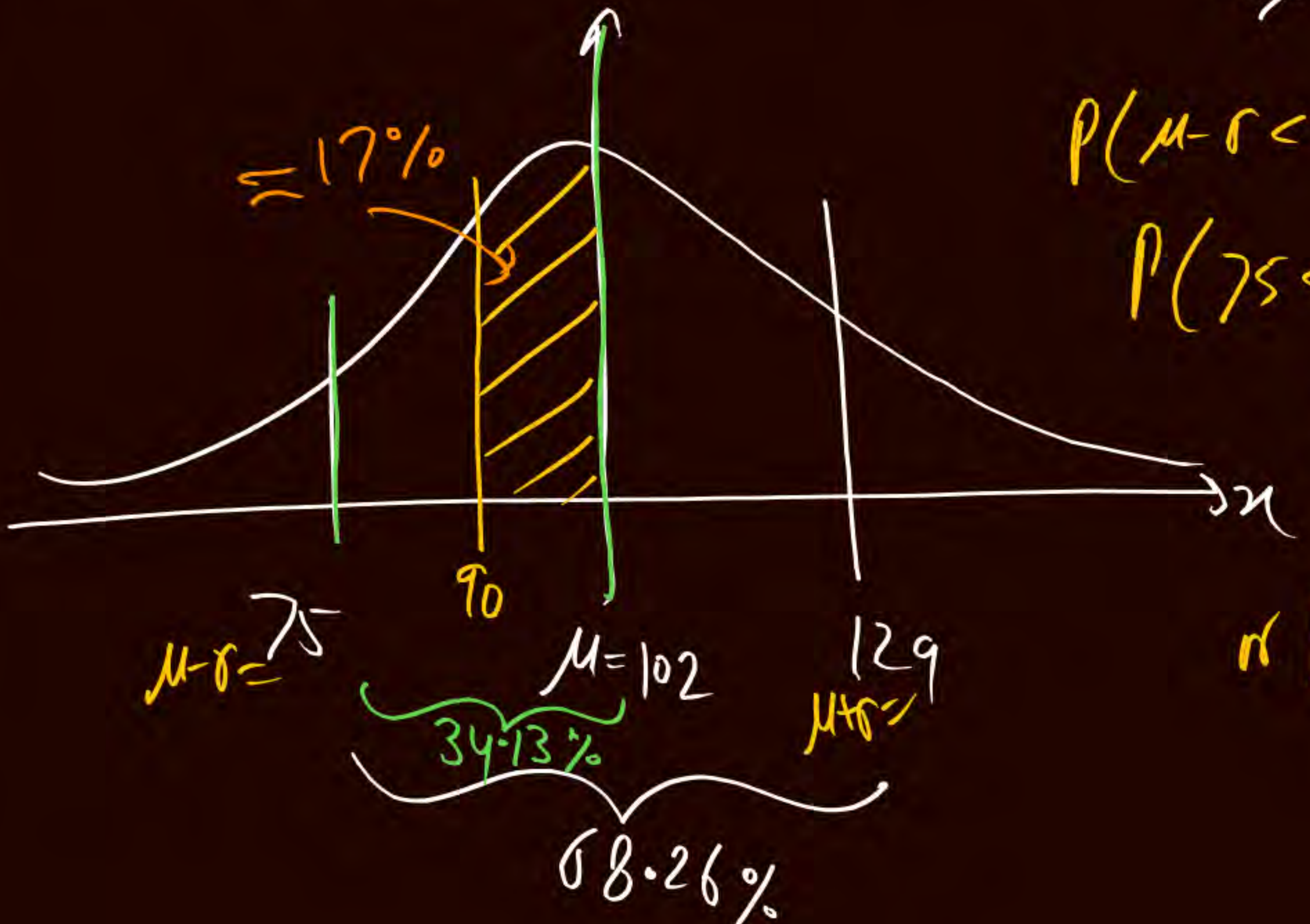
(c) 34%

(d) 16.7%

$x \sim N\{\mu, \sigma^2\} \Rightarrow \mu = 102,\ \sigma = 27$

$\mu - \sigma = 102 - 27 = 75$

$\mu + \sigma = 102 + 27 = 129$



$P(\mu - \sigma < x < \mu + \sigma) = 0.6826$

$P(75 < x < 129) = 0.6826$

$P(75 < x < 102) = 0.3413$

or $P(90 < x < 102) \simeq 17\%$

Q: 2000 students are appearing in an examination in which distribution of marks is assumed to be Normal with average marks $\mu = 30$ & SD $\sigma = 6.25$ marks. Then find the No. of students getting marks b/n 20 & 40. It is given that area under the N. Curve from $z = 0$ to 1.6 is 0.4452?

Sol: $N = 2000$ students,

For single students:

$x = \{\text{Marks obtained}\}$

$\mu = 30$, $\sigma = 6.25$

Note: Large No. of Students are justifying the defⁿ of N. Dist.

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

$$z_1 = \frac{20 - 30}{6.25} = -1.6$$

$$z_2 = \frac{40 - 30}{6.25} = +1.6$$

$$P(20 < x < 40) = P(-1.6 < z < 1.6) = 2P(0 < z < 1.6)$$

$$= 2(0.4452) = 0.8904 = \frac{0.8904}{1}$$

$$= \frac{890.4}{1000} = \frac{1780.8}{2000} \simeq \frac{1781}{2000}$$

So Req Ans = 1781

Q: In a school of 1000 students, Mean Height of student is $\mu = 68.22$ inches and Variance is $\sigma^2 = 10.8$ inches^2, then find the number of students in a school to be over 6 feet tall? It is given that we have 37.46% confidence limits b/w $z = 0$ & 1.15?

Sol: $N = 1000$ students,

For single student:

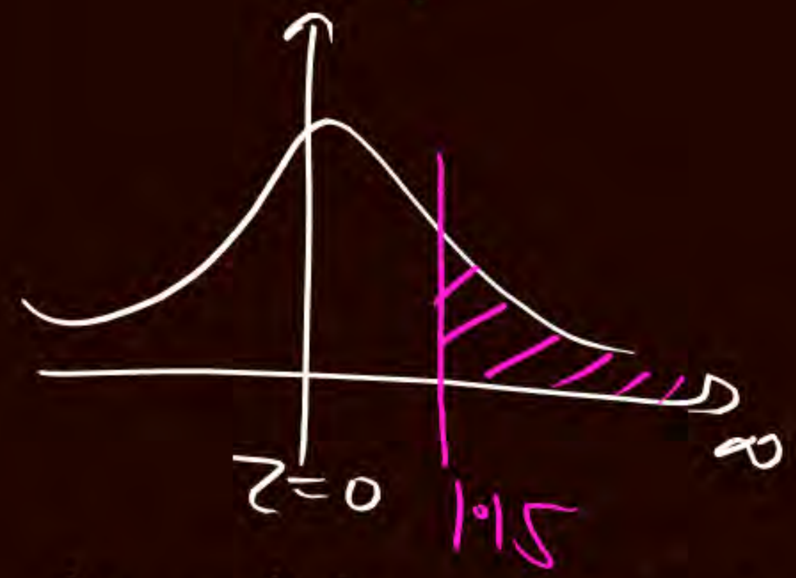
$x = \{\text{Height}\}$

$\mu = 68.22$ inches

$\sigma = \sqrt{10.8}$ "

$$z = \frac{x-\mu}{\sigma} = \frac{72 - 68.22}{\sqrt{10.8}} = 1.15$$

$$P(x > 6) = P(x > 72) = P(z > 1.15) = \frac{1}{2} - P(0 < z < 1.15)$$



$$= 0.5 - 0.3746 = 0.1254$$

$$= \frac{0.1254}{1} = \frac{125.4}{1000} \simeq \frac{125}{1000}$$

i.e out of 1000 students, No. of students having Height > 6 feet = 125

Q: x is Gaussian Variable with mean 1 & Variance 4

& y is also " " " mean -1, & Variance unknown

Also it is given that $\boxed{P(x \le -1) = P(y \ge 2)}$ then find S.D of y?

Sol: $x \sim N\{1, 4\} \Rightarrow Z_x = \dfrac{x - \mu_x}{\sigma_x} = \dfrac{-1 - 1}{2} = -1$

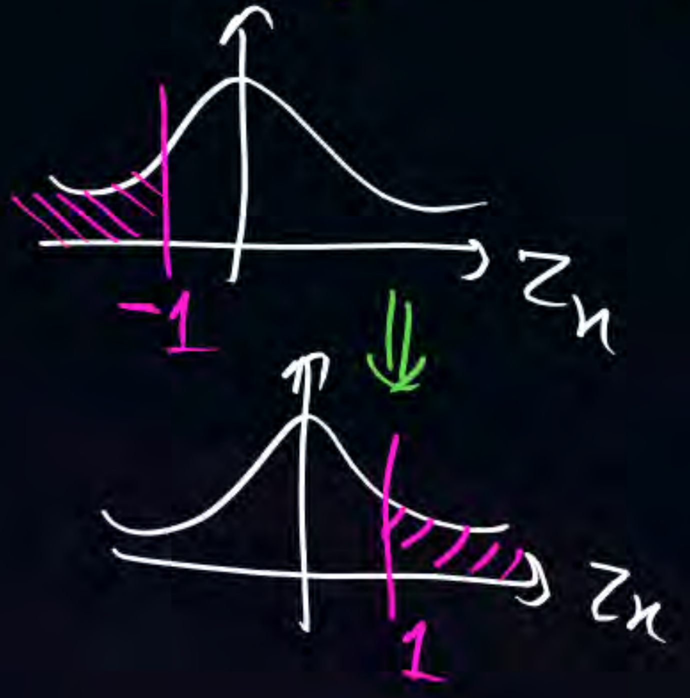
& $y \sim N\{-1, \sigma_y^2\} \Rightarrow Z_y = \dfrac{y - \mu_y}{\sigma_y} = \dfrac{2 - (-1)}{\sigma_y} = \dfrac{3}{\sigma_y}$

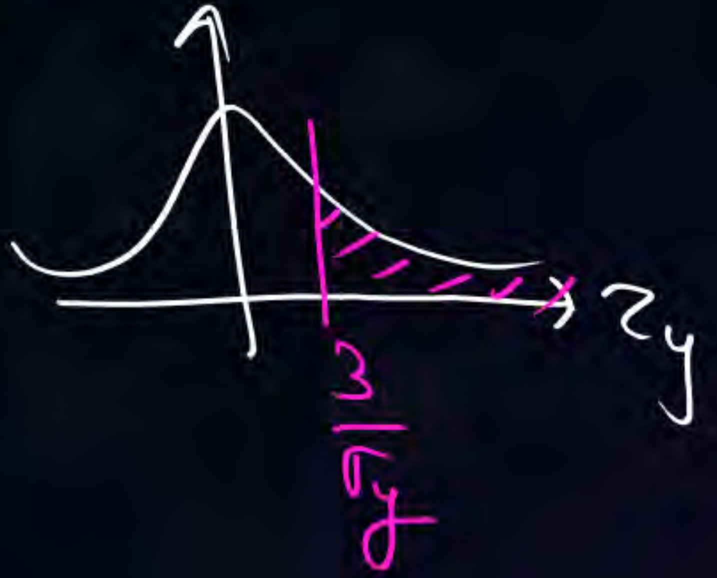
ATQ; $P(x \le -1) = P(y \ge 2)$

$P(Z_x \le -1) = P\left(Z_y \ge \dfrac{3}{\sigma_y}\right)$

$P(Z_x \ge 1) = P\left(Z_y \ge \dfrac{3}{\sigma_y}\right)$

$\Rightarrow 1 = \dfrac{3}{\sigma_y} \Rightarrow \boxed{\sigma_y = 3}$





Wrong App →

$P(x \le -1) = P(y \ge 2)$

step 3

$P(x \ge 1) = P(y \ge 2)$

step 2

$P(Z_x \ge 0) = P\left(Z_y \ge \frac{3}{\sigma_y}\right)$

$\Rightarrow \frac{3}{\sigma_y} = 0$ or $\sigma_y = \infty$ N.D

$\frac{1}{2} = P\left(Z_y \ge \frac{3}{\sigma_y}\right) \Rightarrow \sigma_y = ?$

Q. if x is Zero Mean, Unit Variance Gaussian Variable then find ① $E(x) = ? = 0$

② $E(|x|) = ?$

① Time waste approach :-

$$E(x) = \int_{-\infty}^{\infty} x f(x)\, dx$$

$$= \vdots$$

$$= 0$$

② $x \sim N\{\mu, \sigma^2\}$

$\Rightarrow x \sim N\{0, 1\}$ i.e. x is already in S.N. form and there is no need to convert x into Z & it's p.d.f is

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

$= $ Even fn^n

$$E(|x|) = \int_{-\infty}^{\infty} \underbrace{|x| . f(x)}_{\text{Even func}^n} dx$$

$$= 2\int_0^{\infty} |x| . \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \sqrt{\frac{2}{\pi}} \int_0^{\infty} x . e^{-\left(\frac{x^2}{2}\right)} dx$$

Put $\frac{x^2}{2} = t \Rightarrow x\,dx = dt$

$$= \sqrt{\frac{2}{\pi}} \int_0^{\infty} e^{-t} . dt = \sqrt{\frac{2}{\pi}} \left(\frac{e^{-t}}{-1}\right)_0^{\infty}$$

$$= -\sqrt{\frac{2}{\pi}} \left(e^{-\infty} - 1\right) = \sqrt{\frac{2}{\pi}} = 0.8$$

$$Var(aX+bY) = a^2V(X) + b^2V(Y) + 2ab\,Cov(X,Y)$$

Normal Sum theorem (N.S.T) – If x_1 & x_2 are two Ind Gaussian R. Variable having mean and variance as $\{\mu_1, \sigma_1^2\}$ & $\{\mu_2, \sigma_2^2\}$ resp. then $(x_1 \pm x_2)$ will also be N. R. Variable with

- Mean $= \mu_1 \pm \mu_2$
- Variance $= \sigma_1^2 + \sigma_2^2$

i.e. $x_1 \sim N\{\mu_1, \sigma_1^2\}$ & $x_2 \sim N\{\mu_2, \sigma_2^2\}$ then

- $(x_1 + x_2) \sim N\{\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2\}$
- $(x_1 - x_2) \sim N\{\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2\}$

let $Y = (x_1 - x_2)$

$\rightarrow E(Y) = E(x_1 - x_2) = E(x_1) - E(x_2) = \mu_1 - \mu_2$

$\rightarrow Var(Y) = Var(x_1 - x_2) = Var(x_1) + Var(x_2) - 0 = \sigma_1^2 + \sigma_2^2$

$\mu\{\mu,\sigma^2\}$ $\{\mu,\sigma^2\}$

Q: $x_1 \sim N\{2,3\}$ and $x_2 \sim N\{1,4\}$ s.t x_1 & x_2 are Ind then Random Variable defined by (i) $2x_1+3x_2$ is also ______N.R.V______ ?

(ii) Also find the Nature of the dist. of $x_1 - x_2$? is also N.R.V

Sol: $\mu_1 = 2,\ \sigma_1^2 = 3,\ \mu_2 = 1,\ \sigma_2^2 = 4$

Let $\boxed{X = 2x_1 + 3x_2}$

$$\mu_X = E(X) = E(2x_1+3x_2)$$
$$= 2E(x_1) + 3E(x_2)$$
$$= 2(2) + 3(1)$$
$$= 7$$

$$Var(X) = Var(2x_1+3x_2)$$
$$= 4Var(x_1) + 9Var(x_2) + 2(2)(3)Cov(x_1,x_2)$$
$$= 4(3) + 9(4) + 0 = 12+36 = 48$$

$X \sim N\{7, 48\}$

(ii) $Y = x_1 - x_2$

$Y \sim N\{1, 7\}$

Q: if U & V are two Zero mean Ind Gaussian Variable having variances $\frac{1}{4}$ & $\frac{1}{9}$ resp then find $P(3V \geq 2U) = ?$

Sol: $P(3V \geq 2U) = P(3V - 2U \geq 0) = P(x \geq 0) = P(Z_x \geq 0) = \frac{1}{2}$ Ans

Let $\boxed{x = 3V - 2U}$ $\because$ U & V are Gaussian Variables so will x also (by using N.S.Th)

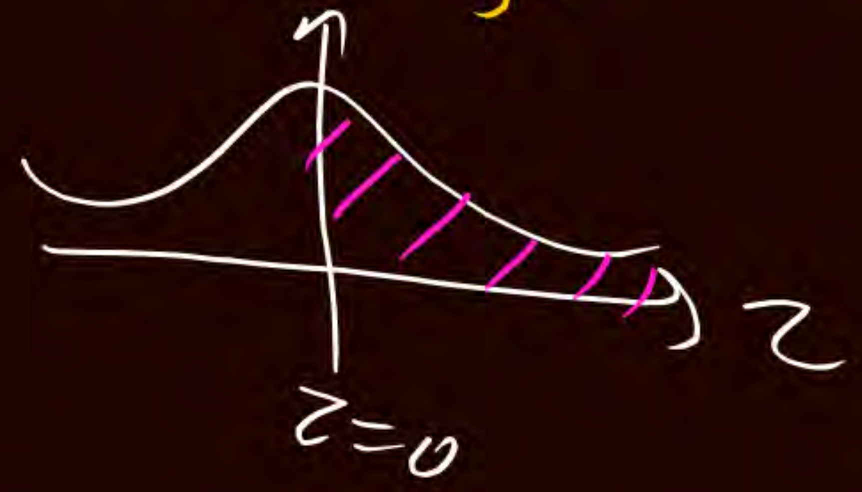
$\mu_x = E(x) = 3E(V) - 2E(U) = 3(0) - 2(0) = 0$

$Var(x) = Var(3V - 2U) = 9Var(V) + 4Var(U) + 0$

$= 9\left(\frac{1}{9}\right) + 4\left(\frac{1}{4}\right) = 2$

is $x \sim N\{0, 2\}$

$Z_x = \frac{x - \mu_x}{\sigma_x} = \frac{0 - 0}{\sqrt{2}} = 0$



Q. Evaluate $\int_a^\infty \frac{1}{b\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-a}{b}\right)^2\right\} dx = ? = \int_\mu^\infty f(x)\,dx = P(\mu < x < \infty)$

$= \frac{1}{2}$

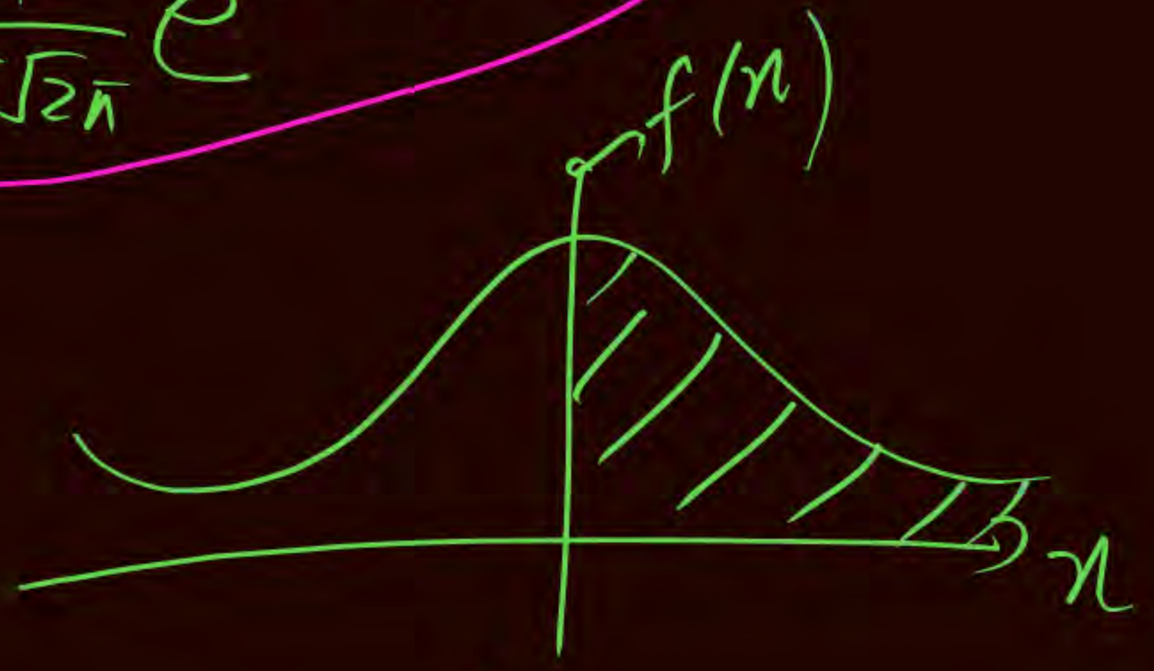
where a & b are statistical Attributes

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

$\Rightarrow x \sim N\{a, b^2\}$

$\Rightarrow a = \mu$

$b = \sigma$



$x = a = \mu$

Correlation :- when two Quantities (x & y) Varies Simultaneously then they are said to be correlated and to their value we will use correlation coeff as;

$$\boxed{r = \frac{Cov(X, Y)}{\sigma_x . \sigma_y}} = \boxed{\frac{E(XY) - E(X).E(Y)}{\sigma_x . \sigma_y}}$$

& it's value lies in b/w $-1 \leq r \leq 1$

1. if $r = 1$ then x & y are perfectly correlated in tve sense
2. if $r = -1$ " " " " " -ve "
3. if $r = 0$ " " are called Ind.

THANK - YOU