Date: 26th March 2021

VIRTUAL COACHING CLASSES
ORGANISED BY BOS (ACADEMIC), ICAI

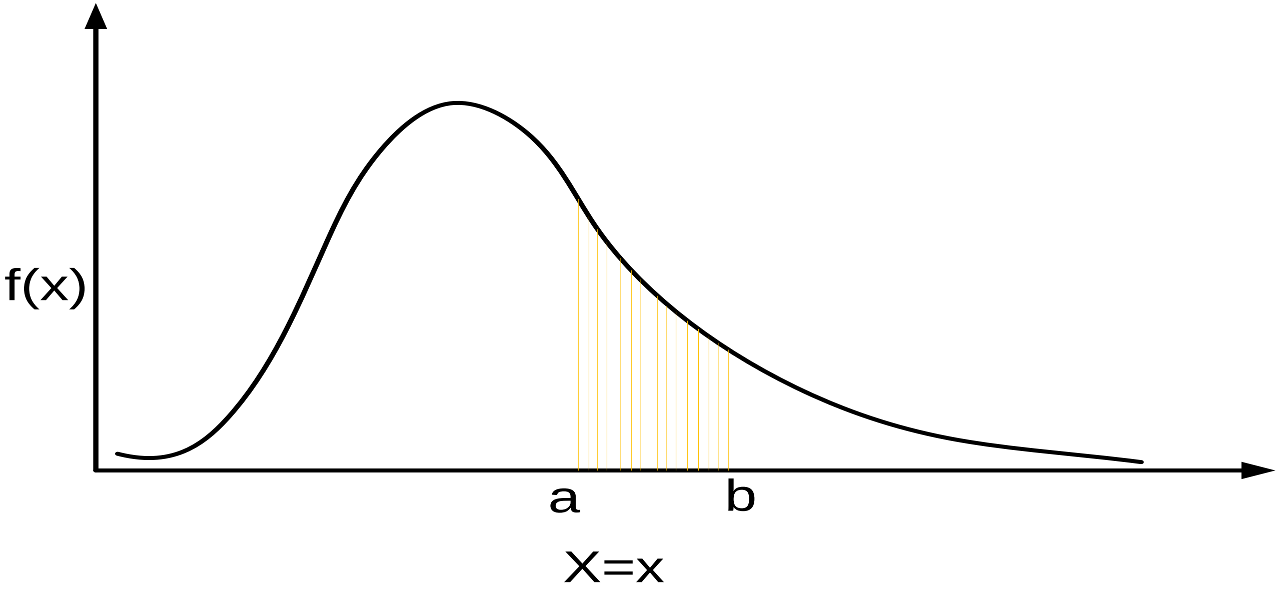
FOUNDATION LEVEL
PAPER 3: BUSINESS MATHEMATICS, LOGICAL REASONING & STATISTICS

Faculty: CA Arijit Chakraborty

Properties of Probability Density Function

he function $f(x)$ is a probability density function for the continuous random ariable X, defined over the set of real numbers R, if

1. $f(x) \geq 0, \text{for all } x \in R$
2. $\int_{-\infty}^{\infty} f(x)dx = 1$
3. $P(a \leq X \leq b) = \int_{a}^{b} f(x)\, dx$
4. $\mu = \int_{-\infty}^{\infty} x f(x)\, dx$
5. $\sigma^{2} = \int_{-\infty}^{\infty} (x-\mu)^{2} f(x) dx$



17.4 NORMAL OR GAUSSIAN DISTRIBUTION

- In case of a continuous random variable like height or weight, **it is impossible to distribute the total probability among different mass points because between any two unequal values, there remains an infinite number of values**.
- Thus a continuous random variable is defined in term of its probability density function f (x), provided, of course, **such a function really exists**, f (x) satisfies the following condition:
- f(x) >= 0 for x belonging to (- infinity, + infinity)

$$F_X(x) = P(X \le x) = \int_{-\infty}^{x} f_X(t)\,dt \qquad \text{for all} \quad x \in R.$$

Important Continuous Distributions

Normal Distribution

A random variable X with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad \text{for} \quad -\infty < x < \infty \qquad (3\text{-}4)$$

has a **normal distribution** (and is called a **normal random variable**) with parameters μ and σ, where $-\infty < \mu < \infty$, and $\sigma > 0$. Also,

$$E(X) = \mu \qquad \text{and} \qquad V(X) = \sigma^2$$

The mean and variance of the normal distribution are derived at the end of this section.

The Normal PDF

It's a probability function, so no matter what the values of μ and σ, must integrate to 1!

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

Normal distribution is defined by its mean and standard dev.

E(X)=μ =

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

Var(X)=σ^2 =

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

Standard Deviation(X)=σ

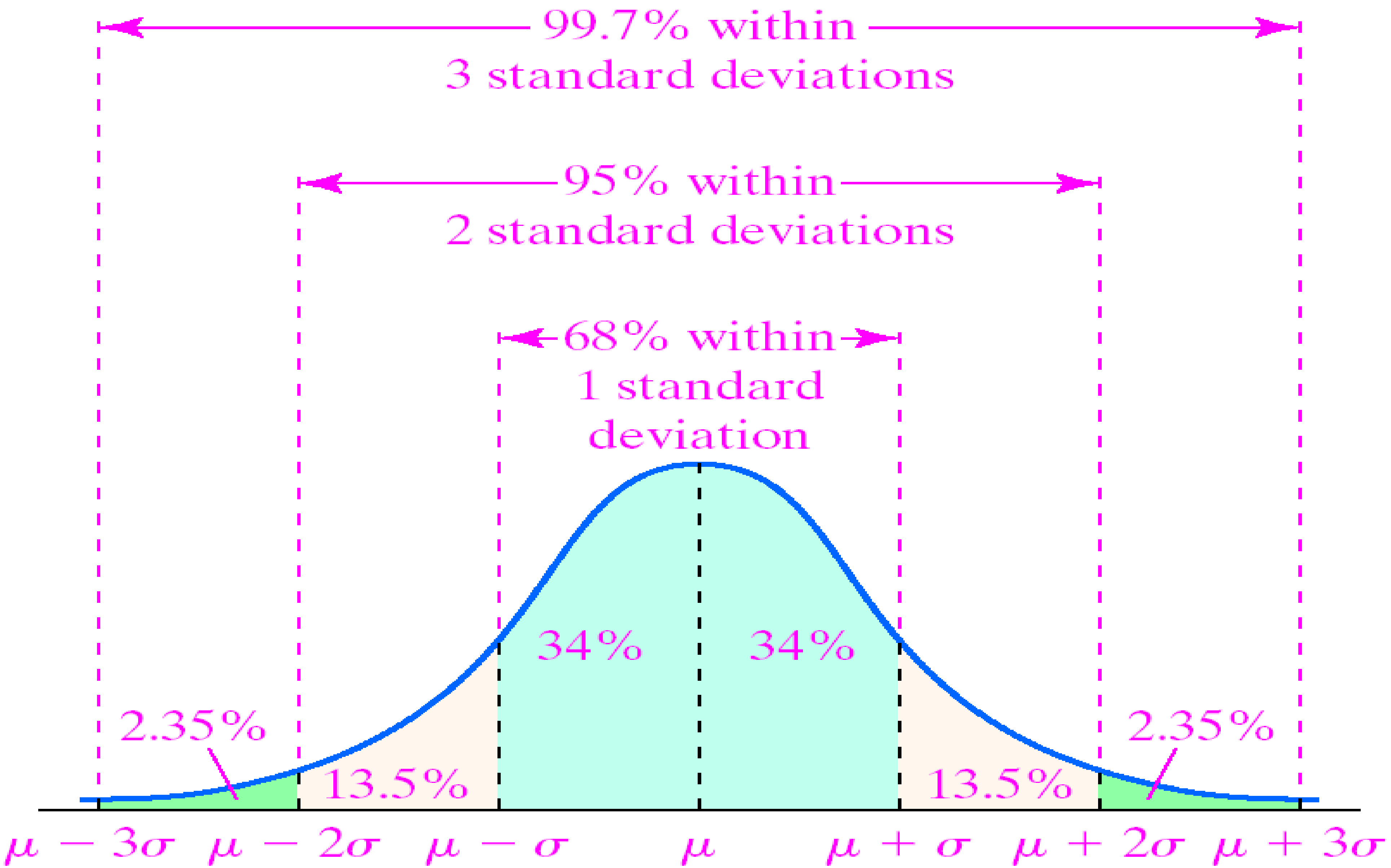
The Normal Distribution – PDF : recap

- A continuous random variable x is defined to follow normal distribution with parameters mew and sigma 2, to be denoted by
- X ~ N (mew, sigma 2) (17.16)
- A random variable *X* is said to have a *normal distribution* if and only if, for σ > 0 and -∞ < μ < ∞, the density function of *X* is

$$f_X(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} \end{cases} \qquad -\infty < x < \infty$$

- The normal distribution is a symmetric distribution and has two parameters μ and σ.
- A very famous normal distribution is the Standard Normal distribution with parameters μ = 0 and σ = 1.

Normal Distribution



Facts and Properties of Pdf

- If X is a continuous random variable with a well-behaved cdf F then

$$f_X(x) = \begin{cases} \frac{d}{dx} F_X(x) & \text{if derivative existes at } x. \\ 0 & \text{otherwize} \end{cases}$$

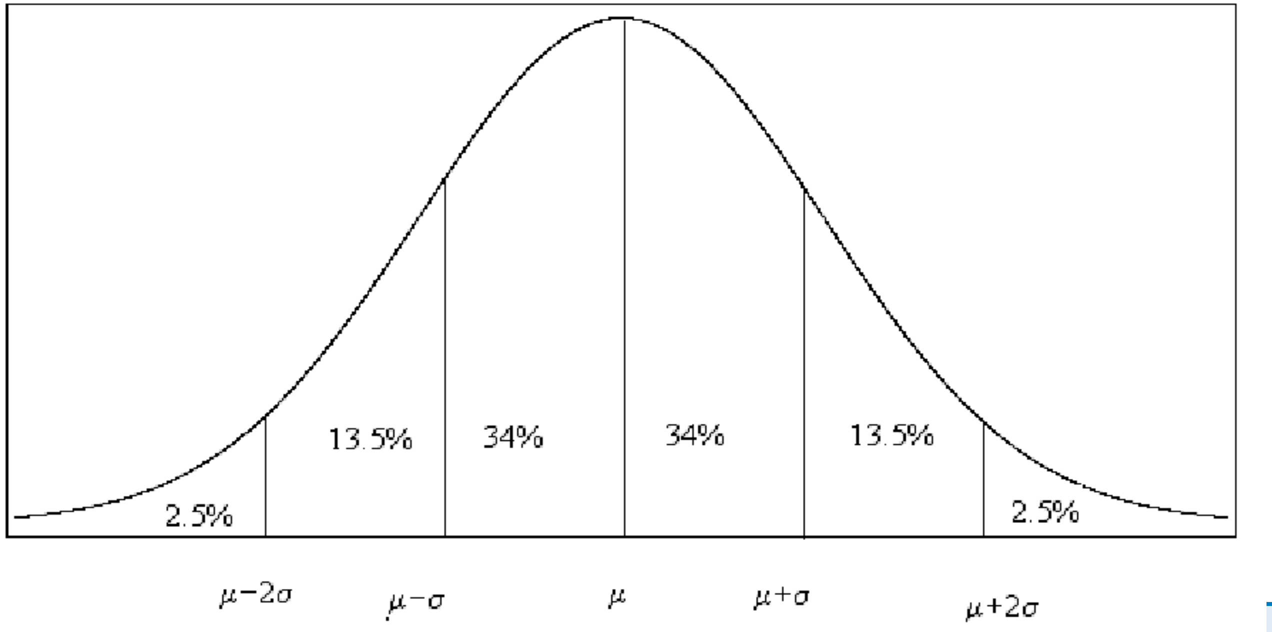
-
  - $f_X(x) \geq 0$ for all $x \in R$)
  - $\int_{-\infty}^{\infty} f_X(x)dx = 1$

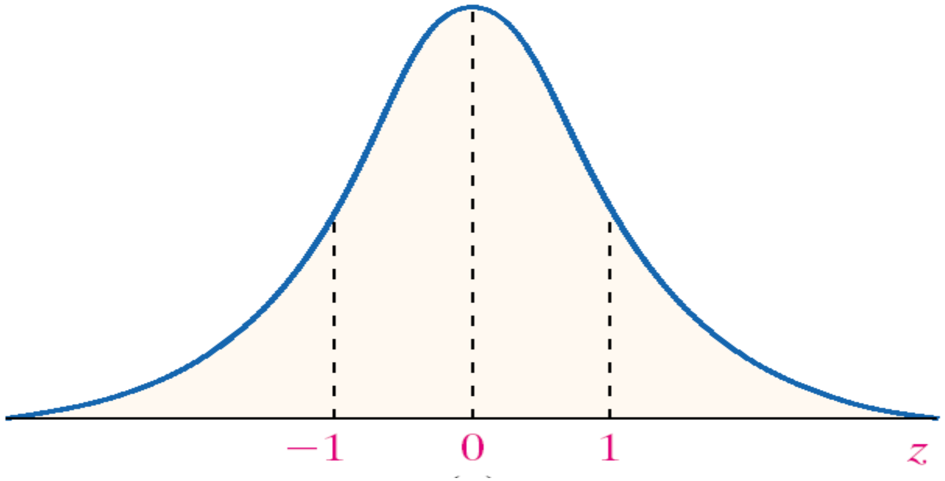
  Any function satisfying these two properties is a probability density function (pdf) for some random variable X.

- **Note:** $f_X(x)$ does **not** give a probability.

- For continuous random variable X with density f
  - $P(a \leq X \leq b) = \int_a^b f_X(t)dt$ why?
  - $P(X = a) = \int_a^a f_X(t)dt = 0$

Standard Normal Distribution





If X is a normal random variable with $E(X) = \mu$ and $V(X) = \sigma^2$, the random variable

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

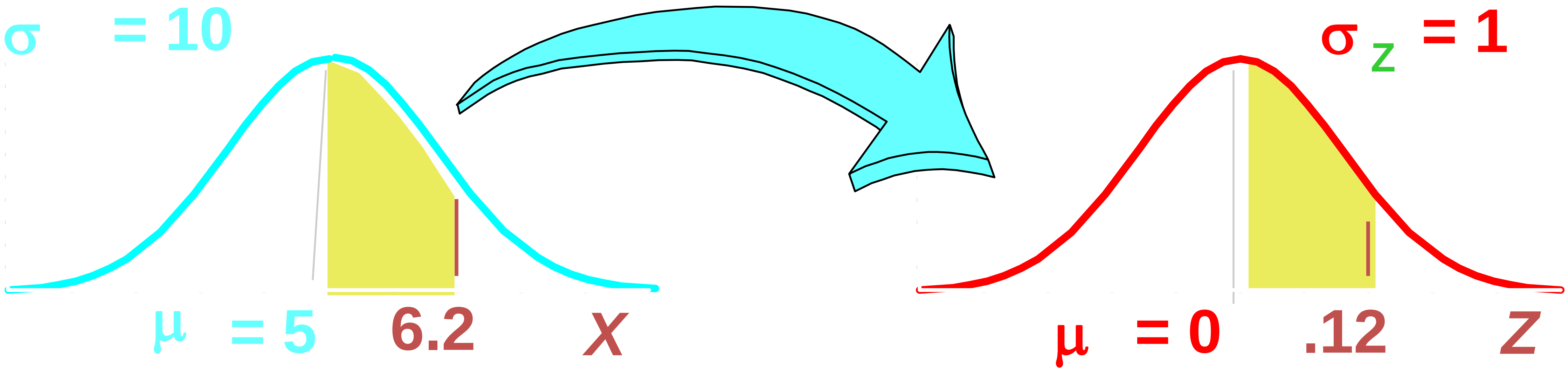
is a normal random variable with $E(Z) = 0$ and $V(Z) = 1$. That is, Z is a **standard normal** random variable.

Standardizing Example

$$Z = \frac{X - \mu}{\sigma} = \frac{6.2 - 5}{10} = 0.12$$

Normal Distribution

Standardized Normal Distribution



Shaded Area Exaggerated

Standard Normal Distribution

The normal distribution has computational complexity to calculate $P(x_1 < x < x_2)$ for any two (x_1, x_2) and given μ and σ

To avoid this difficulty, the concept of z-transformation is followed.

$$\text{z} = \frac{x-\mu}{\sigma} \qquad \text{[Z-transformation]}$$

X: Normal distribution with mean μ and variance σ^2.

Z: Standard normal distribution with mean $\mu = 0$ and variance $\sigma^2 = 1$.

Therefore, if *f(x)* assumes a value, then the corresponding value of $f(z)$ is given by

$$(x\colon \mu, \sigma)\colon P(x_1 < x < x_2) = \frac{1}{\sigma\sqrt{2\pi}} \int_{x_1}^{x_2} e^{-\frac{1}{2\sigma^2}(x-\mu)^2} dx$$

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

Properties of the Normal Distribution

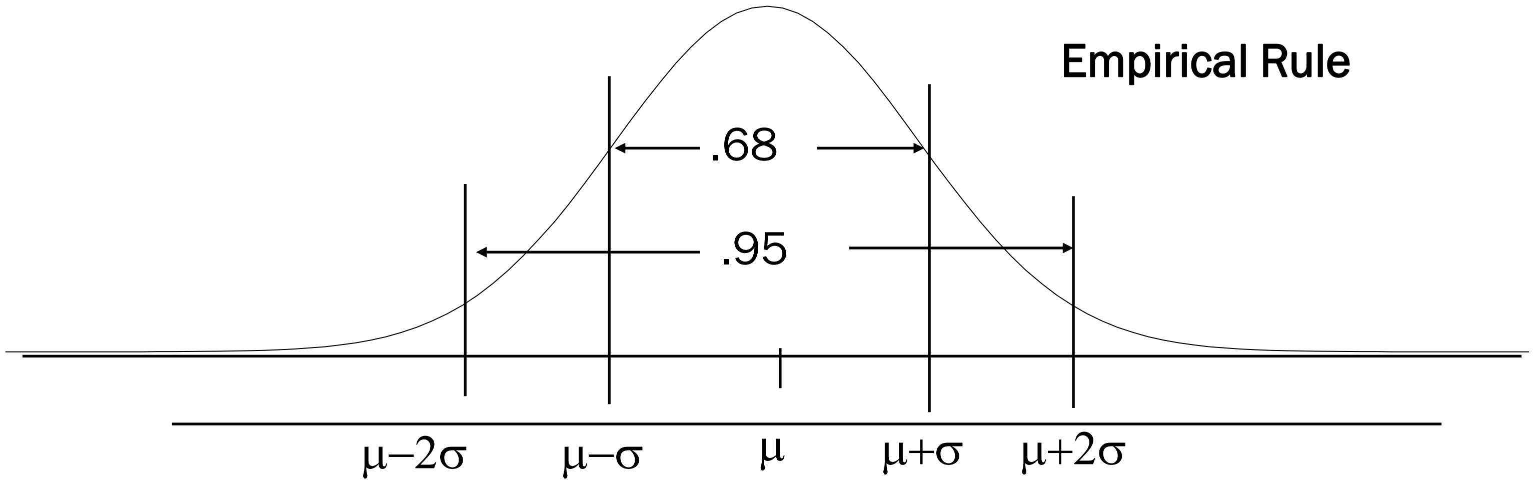
- Symmetric, bell-shaped density function.
- 68% of area under the curve between $\mu \pm \sigma$.
- 95% of area under the curve between $\mu \pm 2\sigma$.
- 99.7% of area under the curve between $\mu \pm 3\sigma$.

$$Y \sim N(\mu, \sigma)$$

$$Y - \mu \sim N(0, \sigma)$$

$$\frac{Y - \mu}{\sigma} \sim N(0,1)$$

Standard Normal Form



Area under Normal curve

- From this figure, we find that
- P (m – s < x < m) = P (m < x < m + s) = 0.34135
- or alternatively, **P (–1 < z < 0) = P (0 < z < 1) = 0.34135** P (m – 2 s < x < m) = P (m < x < m + 2 s) = 0.47725
- i.e. **P (– 2 < z < 1) = P (1 < z < 2) = 0.47725**
- P (m – 3 s < x < m) = P (m < x < m + 3s) = 0.49865 i.e. **P(–3 < z < 0) = P (0 < z < 3) = 0.49865** ……. (17.32)

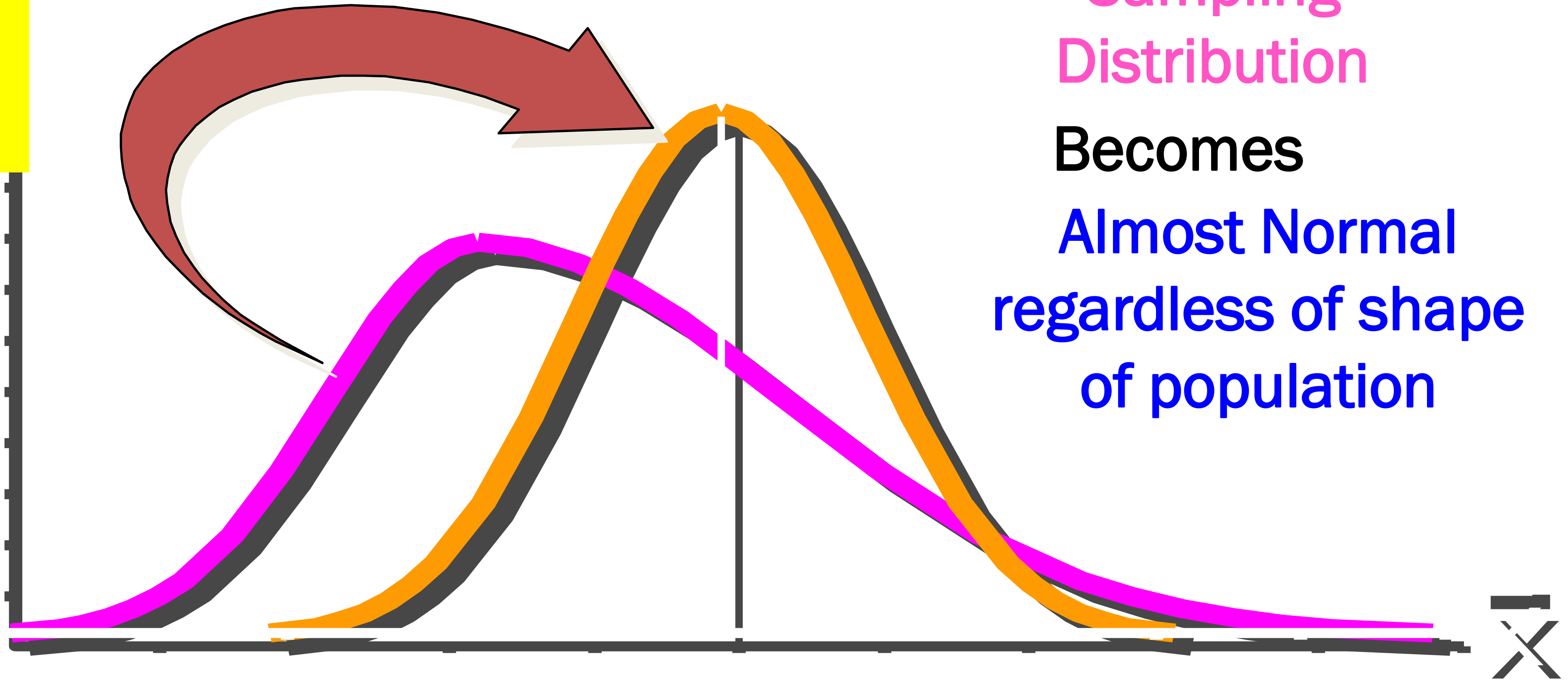
- combining these results, we have P (m – s < x < m + s) = 0.6828
- **=> P (–1 < z < 1) = 0.6828**
- P (m – 2 s < x < m + 2s) = 0.9546
- => **P (– 2 < z < 2) = 0.9546**
- and P (m – 3 s < x < m + 3 s) = 0.9973
- => **P (– 3 < z < 3) = 0.9973.** .... (17.33)
- We note that 99.73 per cent of the values of a normal variable lies between (m – 3 s) and (m + 3 s).
- Thus **the probability that a value of x lies outside that limit is as low as 0.0027.**

Normal approximation

- 1.If the **variable under study does not follow normal distribution**, a simple transformation of the variable, in many a case, would lead to the normal distribution of the changed variable.
- 2. When **n, the number of trials of a binomial distribution, is large** and **p, the probability of a success, is moderate i.e. neither too large nor too small** then the binomial distribution, also, tends to normal distribution.
- **3. Poisson distribution, also for large value of m approaches normal distribution**.

Central Limit Theorem

As Sample Size Gets Large Enough



Sampling Distribution Becomes Almost Normal regardless of shape of population

Normal : using normal distribution table

Example 1

Let $X \sim N(3, 16)$, what is $P(X > 0)$?

$$P(X > 0) = P\left(\frac{X-3}{4} > \frac{0-3}{4}\right) = P\left(Z > -\frac{3}{4}\right) = 1 - P\left(Z \le -\frac{3}{4}\right)$$

$$= 1 - \Phi(-\frac{3}{4}) = 1 - (1 - \Phi(\frac{3}{4})) = \Phi(\frac{3}{4}) = 0.7734$$

What is $P(2 < X < 5)$?

$$P(2 < X < 5) = P\left(\frac{2-3}{4} < \frac{X-3}{4} < \frac{5-3}{4}\right) = P\left(-\frac{1}{4} < Z < \frac{2}{4}\right)$$

$$= \Phi(\frac{2}{4}) - \Phi(-\frac{1}{4}) = \Phi(\frac{1}{2}) - (1 - \Phi(\frac{1}{4})) = 0.2902$$

THANK YOU