Mine Automation and Data Analytics

SWAYAM NPTEL COURSE ON

MINE AUTOMATION AND DATA ANALYTICS

By

Prof. Radhakanta Koner

Department of Mining Engineering

Indian Institute of Technology (Indian School of Mines) Dhanbad

Module 8: Inferential Statistics

Lecture 20 A: Continuous Random Variable Part I

CONCEPTS COVERED

1. Continuous random variable
2. Probability Distribution Function
3. Uniform distribution
4. Exponential distribution
5. Normal distribution

Discrete and Continuous Random Variables

Definition:

A discrete random variable is characterized by its ability to assume, at most, a countable number of possible values.

Consequently, any random variable capable of adopting either a finite number or a countably infinite number of distinct values qualifies as a discrete random variable.

It's worth noting that there are also random variables whose set of potential values is uncountably infinite.

Definition:

Continuous random variables pertain to scenarios where outcomes of random events are numerical, yet cannot be enumerated and are infinitely divisible.

Discrete Random Variable

- A discrete random variable is characterized by having possible values that are distinct points along the real number line.
- Discrete random variables are often associated with counting scenarios

Continuous Random Variable

- A continuous random variable is defined by its possible values spanning an interval along the real number line.
- Continuous random variables typically involve measurement scenarios.

Probability density function (pdf)

- A Probability Density Function (pdf) is a mathematical function that describes the likelihood of a continuous random variable falling within a particular range of values.
- In other words, it provides a way to represent the probability distribution of a continuous random variable. It is denoted by f(x)

Probability density function (pdf)

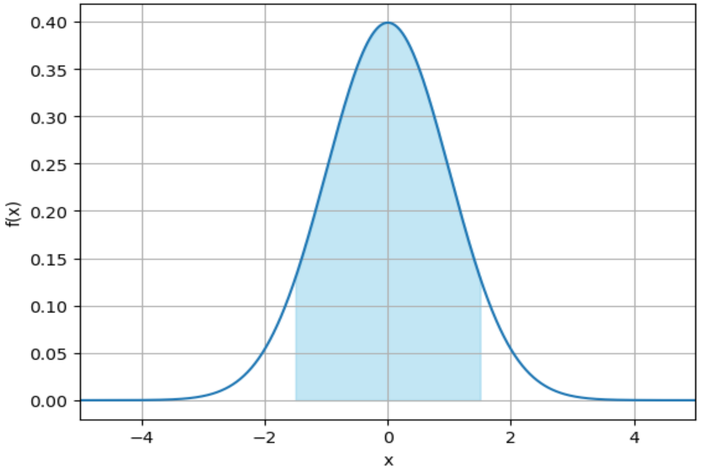
- **The integral of the Probability Density Function curve signifies the likelihood of the random variable existing within a specific interval.**
- **In mathematical terms, for a continuous random variable X with Probability Density Function (f(x)),**

The probability of X lying within the range [a, b] is calculated by integrating the Probability Density Function over that interval.

$$P(X \in [a,b]) = P(a \leq X \leq b) \text{ is area under curve between a and b}$$

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

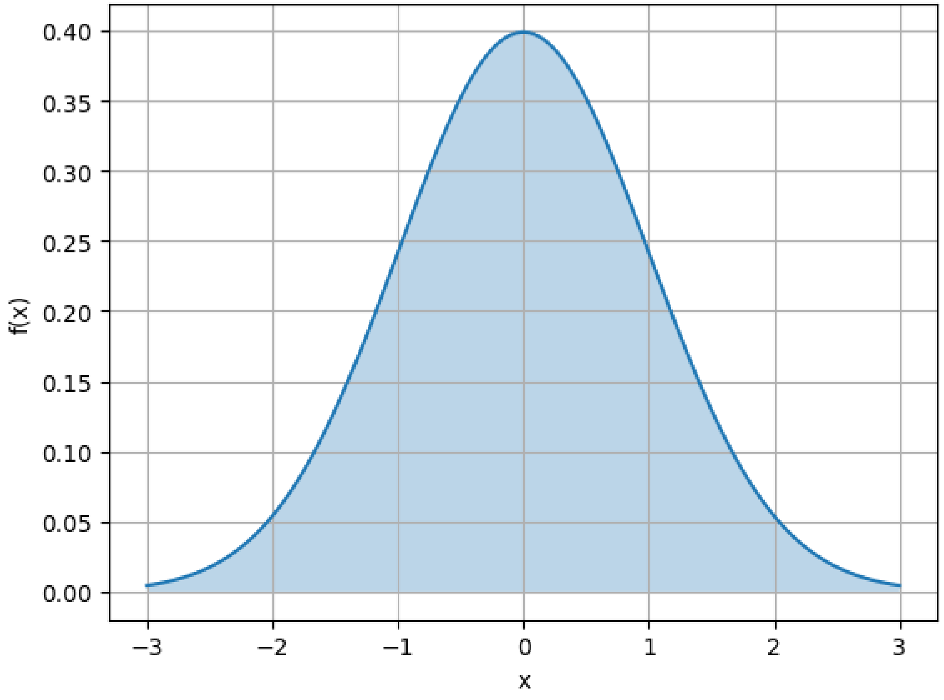
$P(-1.5 \leq X \leq 1.5):$



Properties of pdf

- **The total area under the curve equals 1, ensuring the probabilities encompass all potential outcomes.**

$$\int f(x)dx = 1$$



Properties of pdf

- Therefore, the area beneath the probability distribution curve of a continuous random variable between any two points falls within the range of 0 to 1.

$$0 \leq P(X = x_i) \leq 1$$

- The area beneath the probability density function graph between points a and b remains constant, irrespective of whether the endpoints a and b are inclusive or exclusive.

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

Cumulative Distribution Function

- **For a continuous random variable X**

$$F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$$

- **Given that a probability of continuous random variable (X) taking any single value is zero , we consequently have**

$$P(X \leq a) = P(X < a) = \int_{-\infty}^{a} f(x)\,dx$$

Expectation and Variance of any continuous random variable

Expectation E(X) = $\int x\, f(x)\, dx$

Variance Var(X) = $\int \left(x - E(X)\right)^2 f(x) dx$

Expectation and Variance of any discrete random variable

Expectation E(X) = $\sum_{i=1}^{\infty} x_i P(x_i)$

Variance $Var\ (X) = E(X - \mu)^2$

Continuous Random Variable Distributions

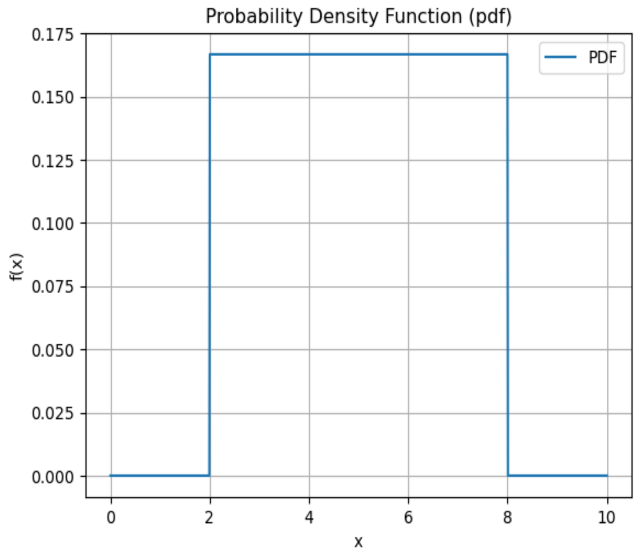
- Uniform Distribution
- Standard Uniform Distribution
- Exponential Distribution
- Normal Distribution
- Standard Normal Distribution
- T distribution
- Chi Squared Distribution

Continuous random variable
Uniform Distribution

A uniform distribution is assigned to a continuous random variable, denoted as X ~ U(a, b).

The Probability Density Function for a uniform distribution is expressed as : $f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & otherwise \end{cases}$

We know that the total area under the probability density curve of any continuous distribution equals 1



$$\int_a^b f(x)dx = 1$$

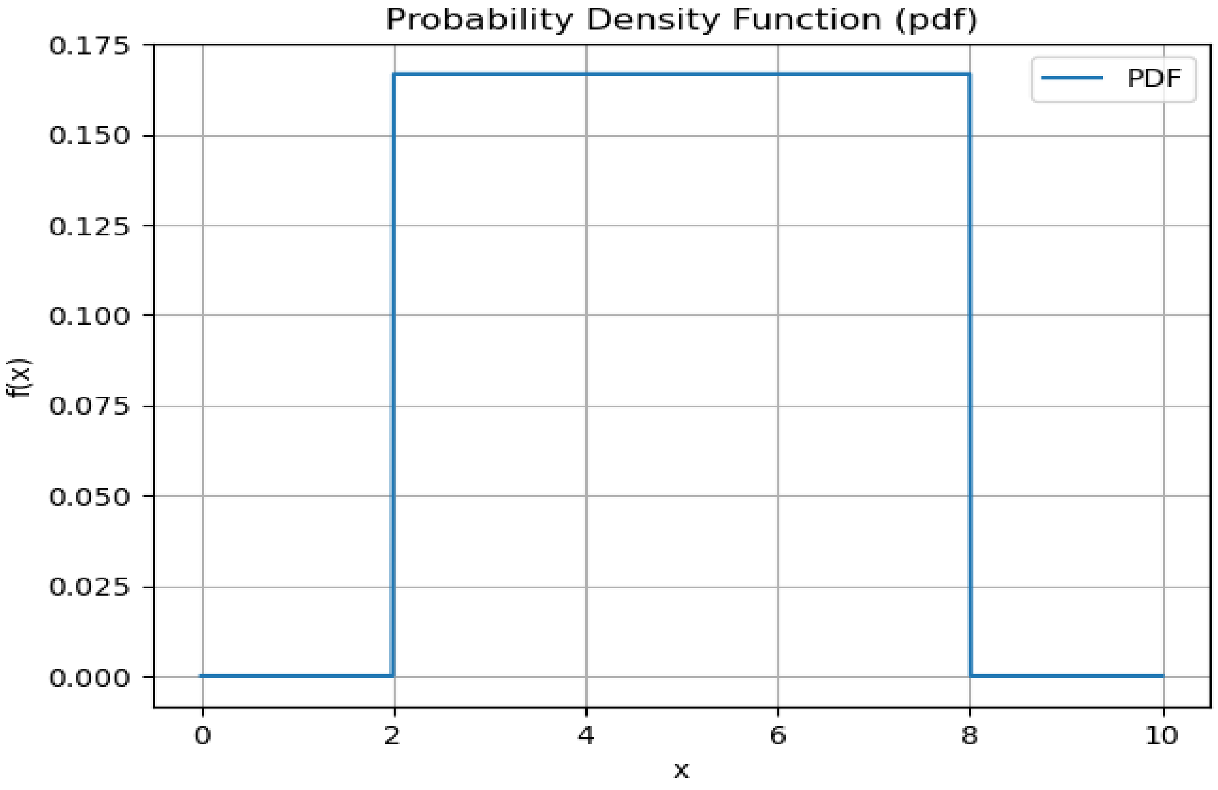
$$f(x) \int_a^b dx = 1$$

$$f(x)\ (b - a) = 1$$

$$f(x) = \frac{1}{b-a}$$

Continuous random variable
Uniform Distribution

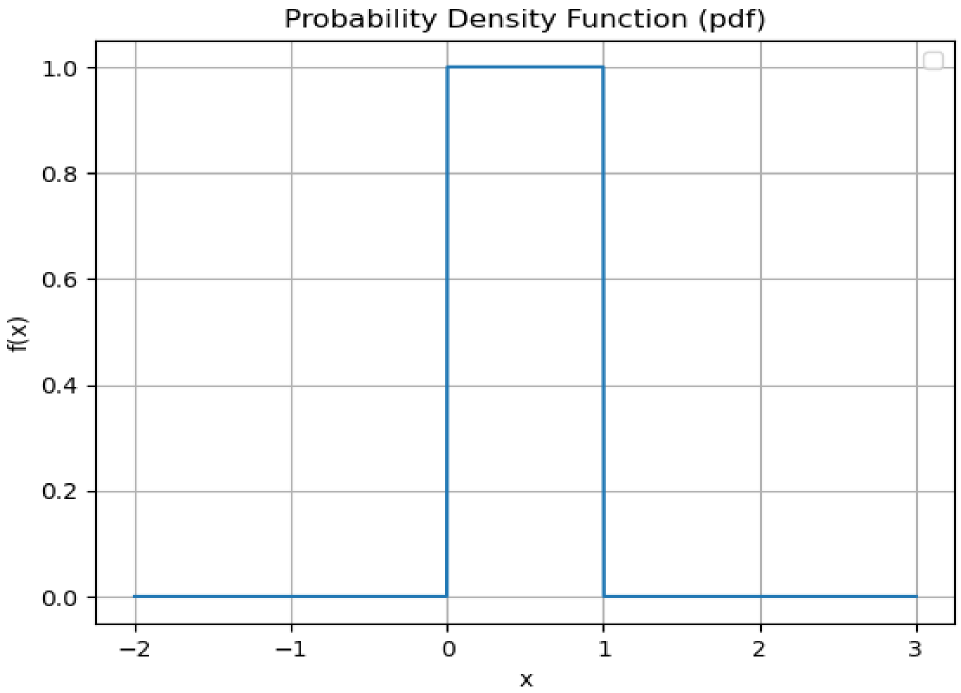
Example: Uniform Distribution U(2,8)



Standard Uniform Distribution

A continuous random variable follows the standard uniform distribution with a minimum value of 0 and a maximum value of 1, represented as X ~ U(0, 1).

Probability Density Function of uniform distribution is given by : $f(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & otherwise \end{cases}$



$$\int_0^1 f(x)dx = 1$$

$$f(x) \int_0^1 dx = 1$$

$$f(x)\ (1 - 0) = 1$$

$$f(x) = 1$$

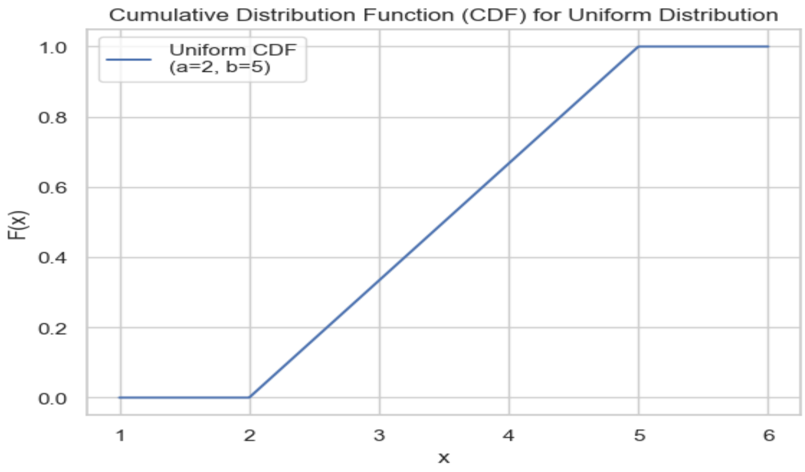
JAN 2024

Cumulative Distribution Function
Uniform Distribution

The cumulative distribution function for a uniform distribution, denoted by F(x), is given by:

$$\boldsymbol{F(x)} = \begin{cases} 0 & x < a \\ \dfrac{x-a}{b-a} & a \le x < b \\ 1 & x \ge b \end{cases}$$

- **We know for any continuous random variable X :** $\boldsymbol{F(a) = P(X \le a) = \int_{-\infty}^{a} f(x)dx}$



$$\boldsymbol{F(x) = P(X \le x) = \int_{-\infty}^{x} f(x)dx}$$

$$\boldsymbol{= \int_{-\infty}^{a} f(x)dx + \int_{a}^{x} f(x)dx}$$

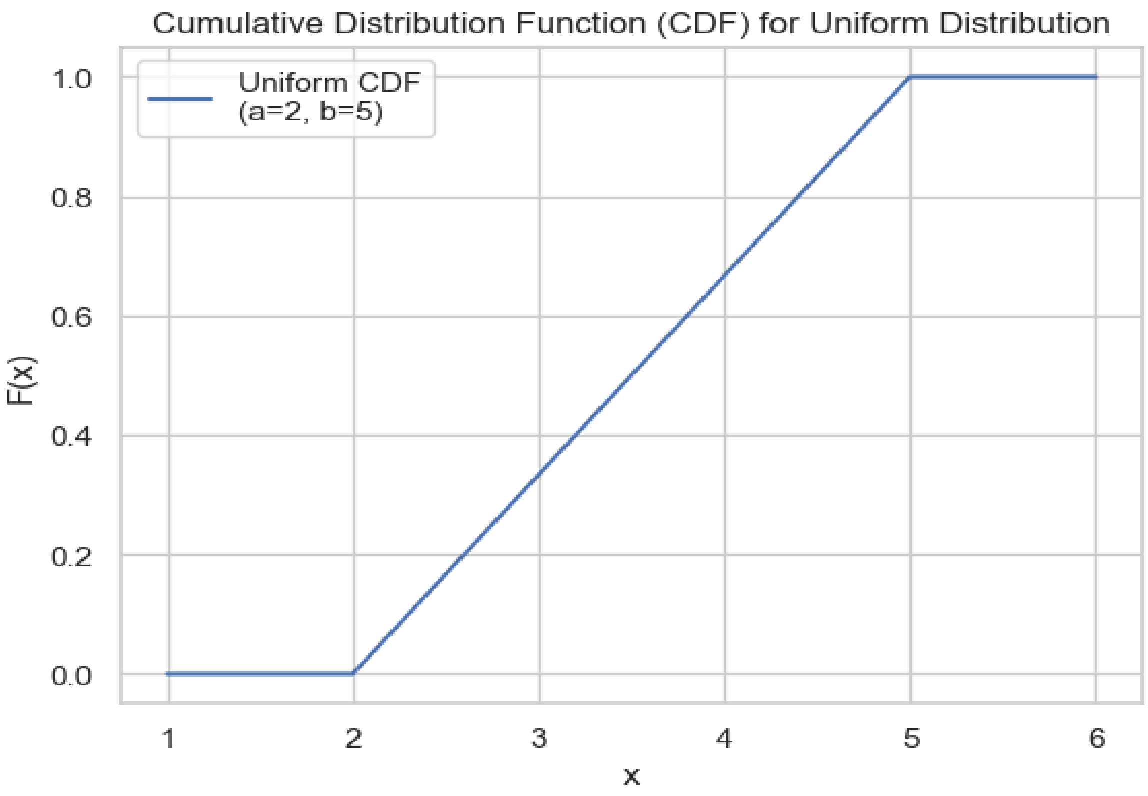
$$\boldsymbol{= 0 + \int_{a}^{x} f(x)dx}$$

$$\boldsymbol{= \int_{a}^{x} \frac{1}{b-a} dx}$$

$$\boldsymbol{= \frac{x-a}{b-a}}$$

Cumulative Distribution Function
Uniform Distribution

Example : cumulative distribution function for uniform distribution U (2,5)

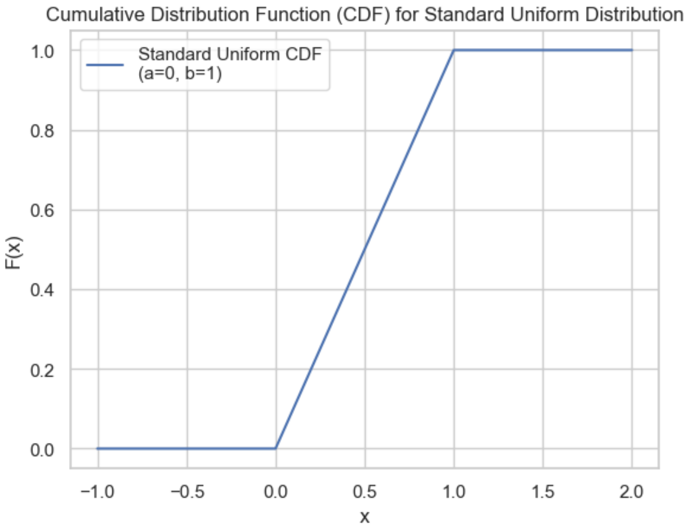


Cumulative Distribution Function
Standard Uniform Distribution

The cumulative distribution function for the standard uniform distribution is denoted by F(x).

$$\boldsymbol{F(x)} = \begin{cases} 0 & x < 0 \\ x & 0 \le x < 1 \\ 1 & x \ge 1 \end{cases}$$

- **We know for any continuous random variable X :** $\boldsymbol{F(a) = P(X \le}$ **a) =** $\boldsymbol{\int_{-\infty}^{a} f(x)dx}$



$\boldsymbol{F(x) = P(X \le}$ **x) =** $\boldsymbol{\int_{-\infty}^{x} f(x)dx}$

= $\boldsymbol{\int_{-\infty}^{0} f(x)dx}$ **+** $\boldsymbol{\int_{0}^{x} f(x)dx}$

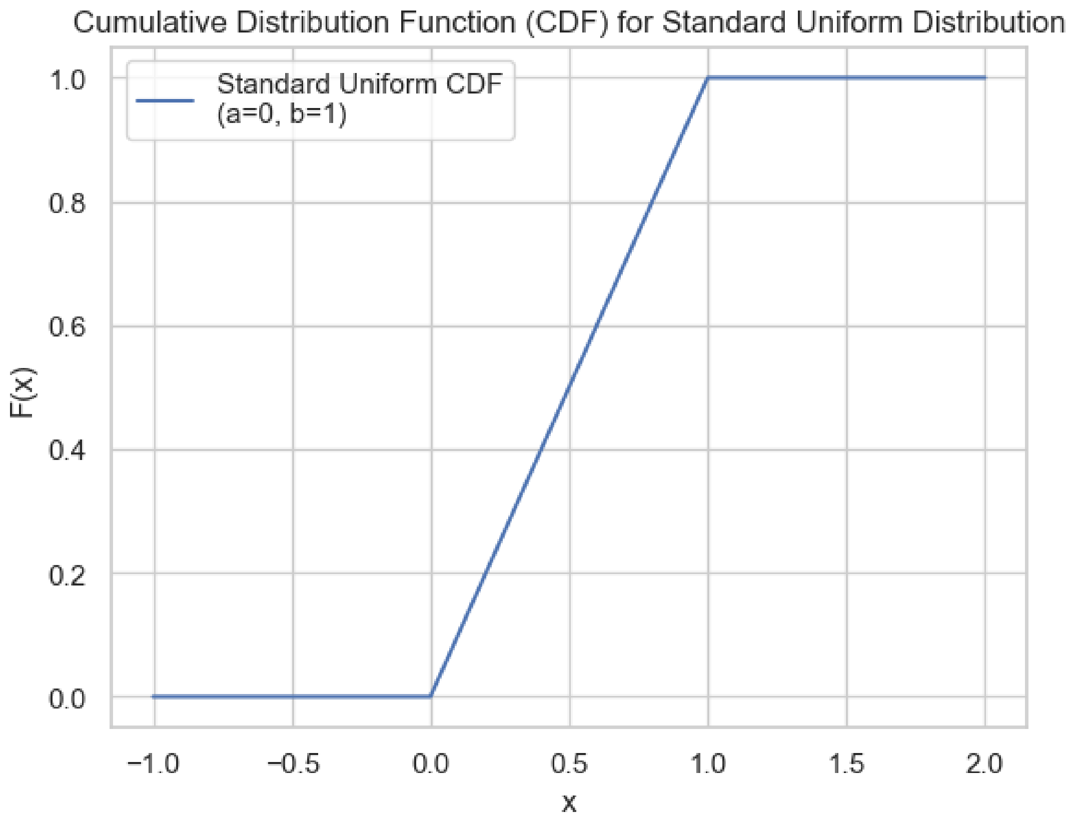
= 0 + $\boldsymbol{\int_{0}^{x} f(x)dx}$

= $\boldsymbol{\int_{0}^{x} 1\, dx}$

= $\boldsymbol{x}$

Cumulative Distribution Function
Standard Uniform Distribution

Example : cumulative distribution function for uniform distribution U (0,1)



Expectation

Expectation of continuous random variable

X ~ U(a, b)

$$E(X) = \frac{a+b}{2}$$

We know $$E(X) = \int_a^b xf(x)dx$$

$$= \int_a^b x\frac{1}{b-a}dx = \frac{1}{b-a}\int_a^b xdx$$

$$= \frac{1}{b-a}\frac{b^2-a^2}{2}$$

Expectation of Uniform Continuous random variable = $\frac{b+a}{2}$

Variance

Variance of continuous random variable

X ~ U(a, b)

$$V(X) = \frac{(b-a)^2}{12}$$

$$V(X) = E\left(X^2\right) - (E(X))^2$$

$$E\left(X^2\right) = \int_a^b x^2 f(x)\,dx$$

$$= \int_a^b x^2 \frac{1}{b-a}\,dx = \frac{1}{b-a}\int_a^b x^2\,dx$$

$$= \frac{1}{b-a}\frac{b^3-a^3}{3}$$

$$= \frac{b^2+a^2+ab}{3}$$

Variance

Variance of continuous random variable

X ~ U(a, b)

$$V(X) = \frac{(b-a)^2}{12}$$

$$V(X) = E\left(X^2\right) - (E(X))^2$$

$$= \frac{b^2+a^2+ab}{3} - \left(\frac{b+a}{2}\right)^2$$

$$= \frac{b^2+a^2+ab}{3} - \frac{a^2+b^2+2ab}{4}$$

$$= \frac{b^2+a^2-2ab}{12}$$

Variance of Uniform Continuous random variable $= \frac{(b-a)^2}{12}$

JAN 2024

Example

X is a uniform random variable over (0,1). Find the probabilities as follows:

1. P(X < 0.5) = 0.5

$P(X \leq x) = P((X \leq x) = P((X < 0.5) = \int_0^{0.5} f(x)dx$

$= \int_0^{0.5} 1.\ dx$ **[We know pdf value for standard uniform distribution is 1]**

= 0.5

1. P(X ≥ 0.8) = 1 – 0.8 = 0.2

$P((X \geq 0.8) = 1 - P((X < 0.8) = 1 - \int_0^{0.8} f(x)dx$

$= 1 - \int_0^{0.8} 1.\ dx$

= 1 - 0.8

= 0.2

Example

X is a uniform random variable over (0,1). Find the probabilities as follows:

1. P(0.2 ≤ X < 0.9) = 0.9 – 0.2 = 0.7

$P(y \leq X \leq$ x) = P (($X \leq$ x) - P (($X \leq$ y) = P (($X \leq$ 0.9) - P (($X \leq$ 0.2)

$$= \int_0^{0.9} f(x)dx - \int_0^{0.2} f(x)dx$$

= 0.9 – 0.2

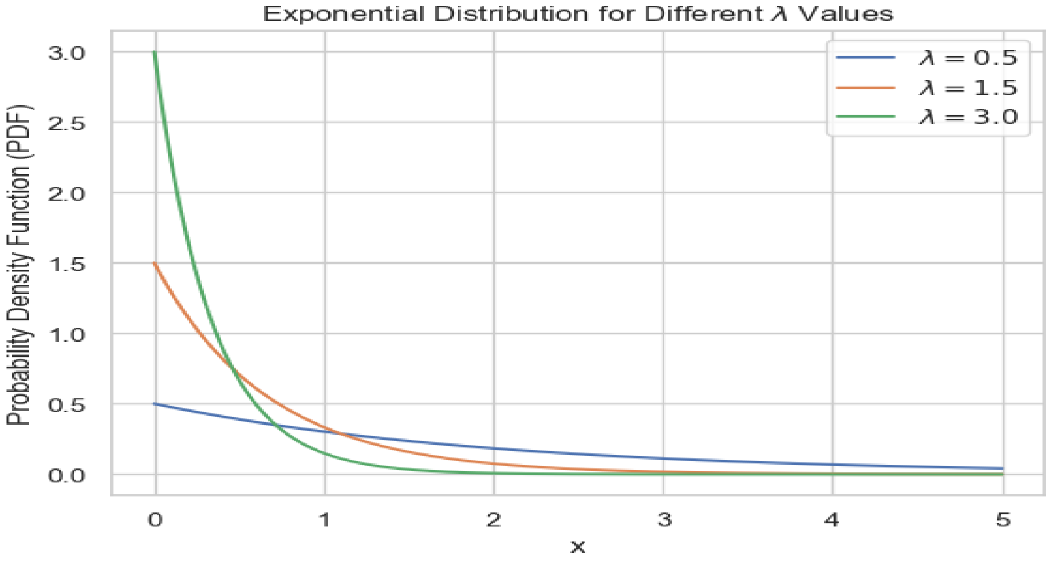
= 0.7

Continuous Random Variable
Exponential Distribution

For some λ > 0, the probability density function (denoted by f(x)) of an exponential continuous random variable is defined as:

$$X \sim Exp(\lambda)$$

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & otherwise \end{cases}$$



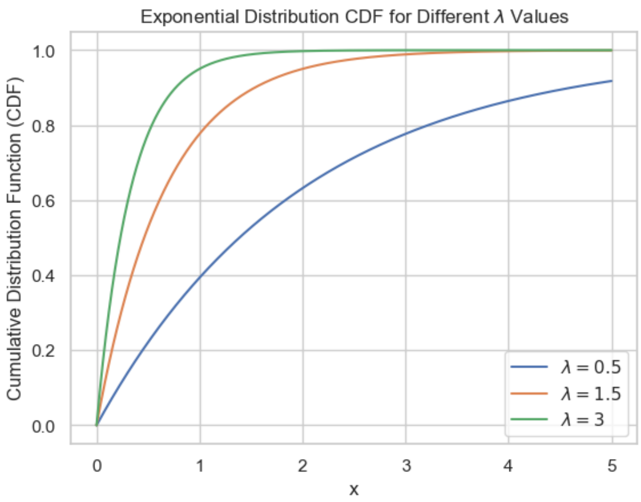
Cumulative Distribution Function
Exponential Distribution

The cumulative distribution function of the exponential distribution is denoted as F(a).

We know cumulative distribution function for any continuous random variable, $F(a) = P(X \leq a) = \int_{-\infty}^{a} f(x)\,dx$

$$= \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{a} f(x)\,dx$$

We Know f(x) = 0 for x less than zero in exponential distribution, So



$$= \int_{0}^{a} \lambda e^{-\lambda x}\,dx$$

$$F(a) = 1 - e^{-\lambda a}$$

Expectation
Exponential Distribution

The Expectation of exponential distribution is given by E(X)

$$E(X^n) = \int_0^\infty x^n \; f(x)dx$$

$$= \int_0^\infty x^n \, \lambda e^{-\lambda x} \, dx$$

$$= \lambda \int_0^\infty x^n \, e^{-\lambda x} \, dx$$

By applying integration by parts formula: $\int udv = uv - \int vdu$

$E(X^n) = \frac{n}{\lambda} E(X^{n-1})$

For n = 1, $E(X^1) = \frac{1}{\lambda} E(X^{1-1}) = \frac{1}{\lambda}$

Expectation of a Exponential Distribution Exp(λ) = $\frac{1}{\lambda}$

Variance
Exponential Distribution

The Variance of exponential distribution is given by V(X)

For n = 1, $\boldsymbol{E(X^1) = \frac{1}{\lambda} E(X^{1-1}) = \frac{1}{\lambda}}$

For n = 2, $\boldsymbol{E(X^2) = \frac{2}{\lambda} E(X^{2-1}) = \frac{2}{\lambda^2}}$

$$\boldsymbol{E(X) = \frac{1}{\lambda}}$$

$$\boldsymbol{V(X) = E(X^2) - (E(X))^2}$$

$$\boldsymbol{= \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2}$$

$$\boldsymbol{= \frac{1}{\lambda^2}}$$

The Variance V(x) of an exponential distribution defined by Exp (λ) = $\boldsymbol{\frac{1}{\lambda^2}}$

Normal Distribution

- The normal distribution, alternatively referred to as the Gaussian distribution or bell curve, is a continuous probability distribution exhibiting symmetry around its mean, which coincides with its median and mode.

- Defined by its mean (μ) and standard deviation (σ), the normal distribution takes on a bell-shaped form.

- Considered paramount in statistics for both theoretical understanding and practical application, the normal distribution holds significant importance.

Probability Density Function (pdf)
Normal Distribution

The continuous random variable X follows a normal distribution, denoted as X ~ N (μ, σ^2) , where μ is the mean and σ^2 is the variance.

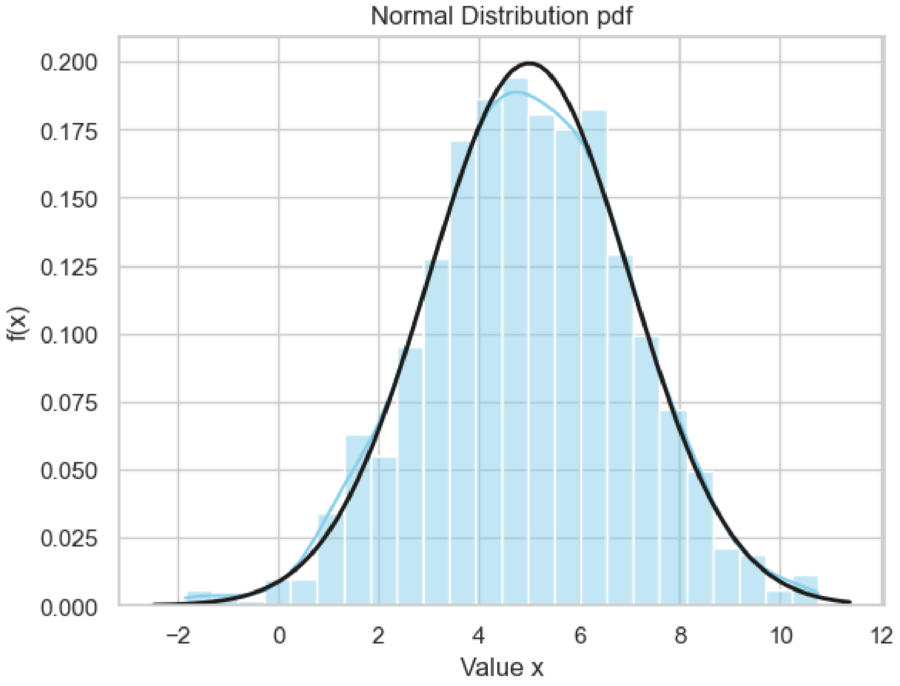
where μ = expected value i.e. mean

σ^2 = variance

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$

Shape of Normal Probability Density Function (pdf)

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$



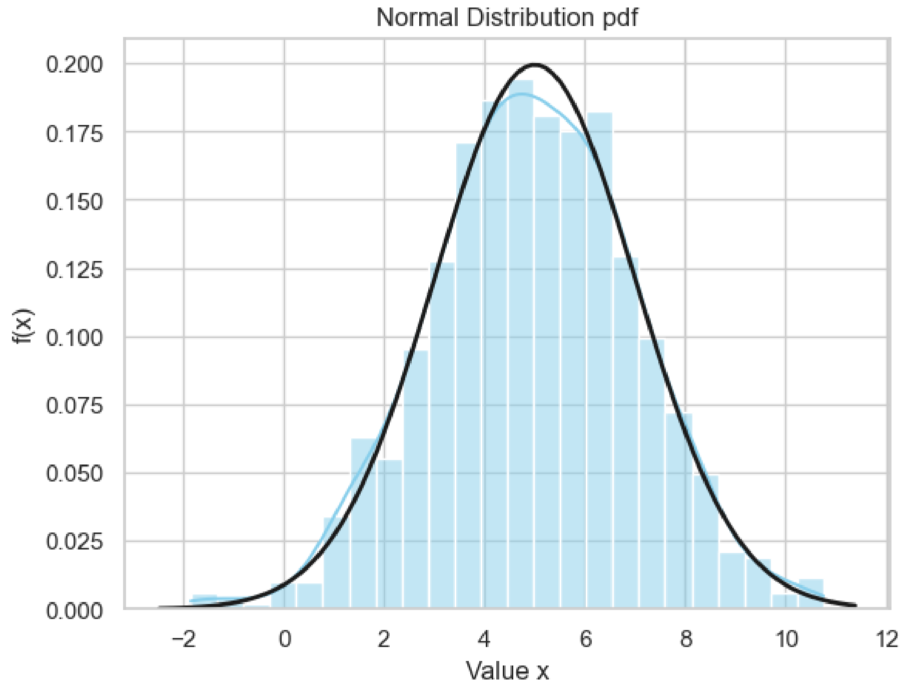
μ = expected value i.e. mean = 5

σ = Standard deviation = 2

Height of the graph

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$

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



$f(x) = 0.199\, for$

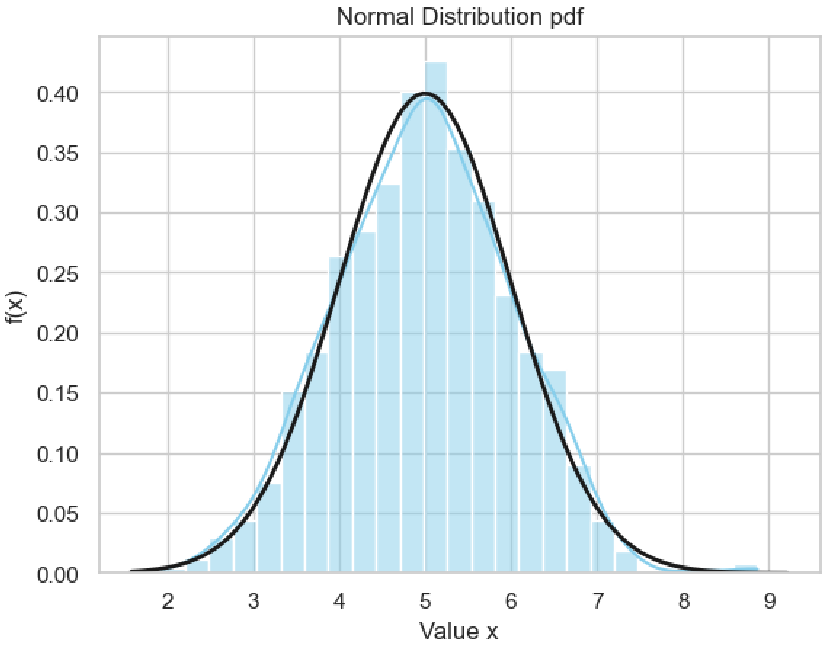
μ = expected value i.e. mean = 5

σ = Standard deviation = 2

Height of the graph

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi\sigma}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$

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



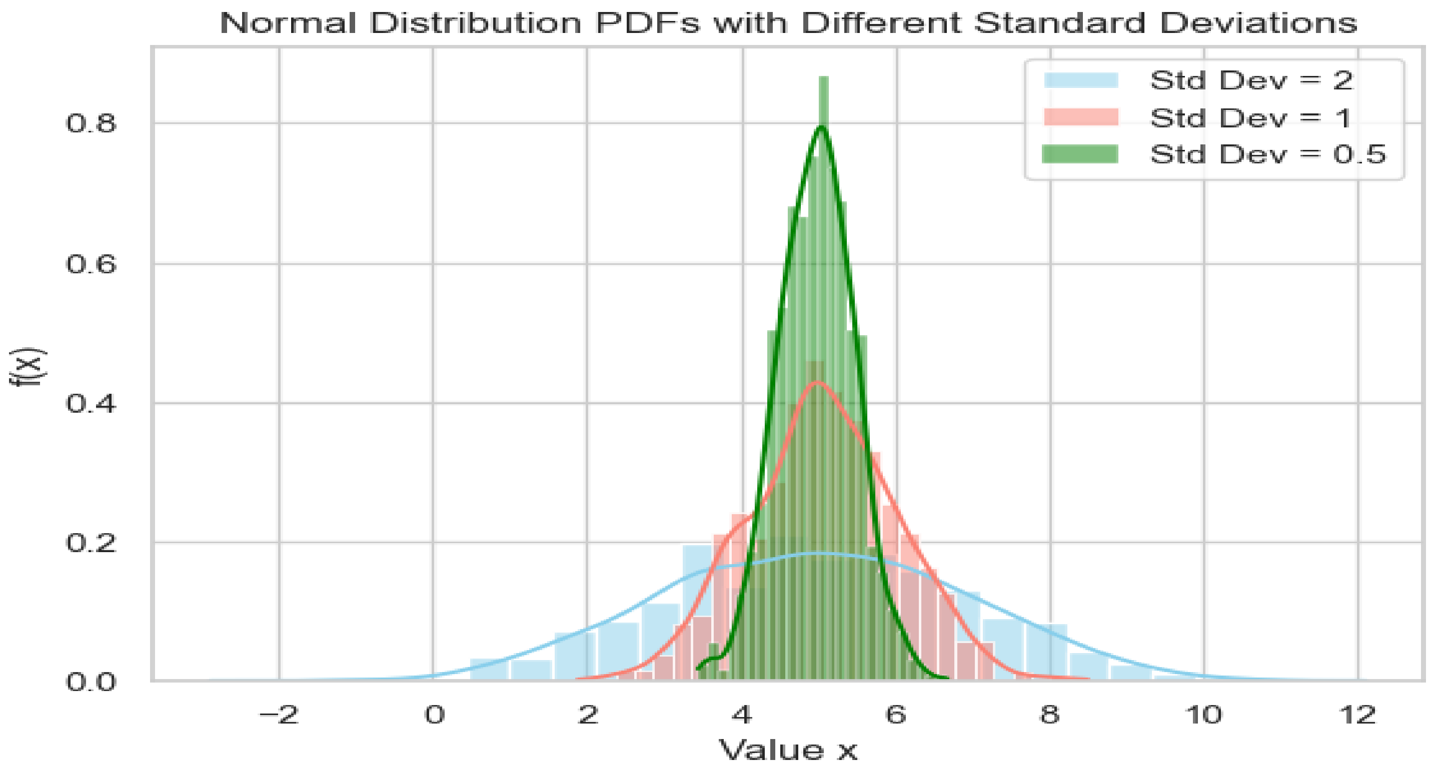
$f(x) = 0.399 \ \ for$

μ = expected value i.e. mean = 5

σ = Standard deviation = 1

Same Mean and Different Standard Deviation

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$



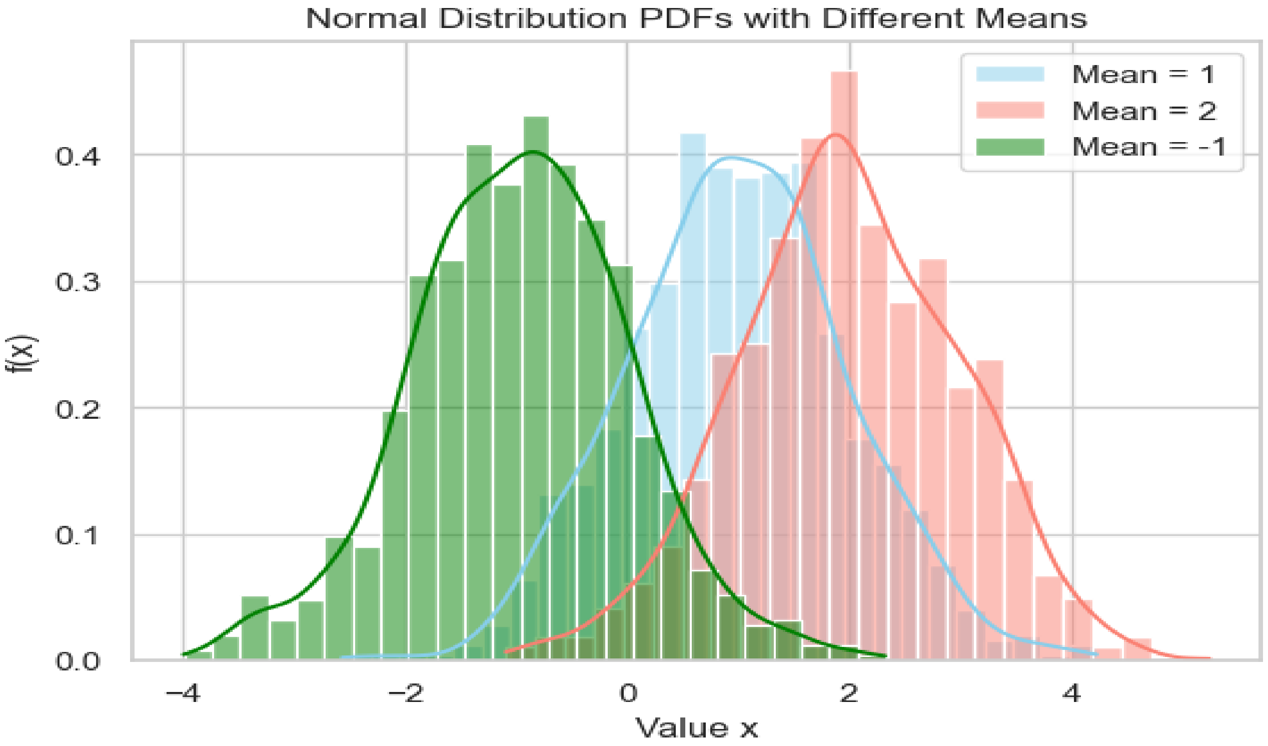
μ = expected value i.e. mean = 5

σ = Standard deviation = 2, 1, 0.5

The larger the standard deviation , the flatter the graph becomes.

Same Standard Deviation and Different Mean

Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$



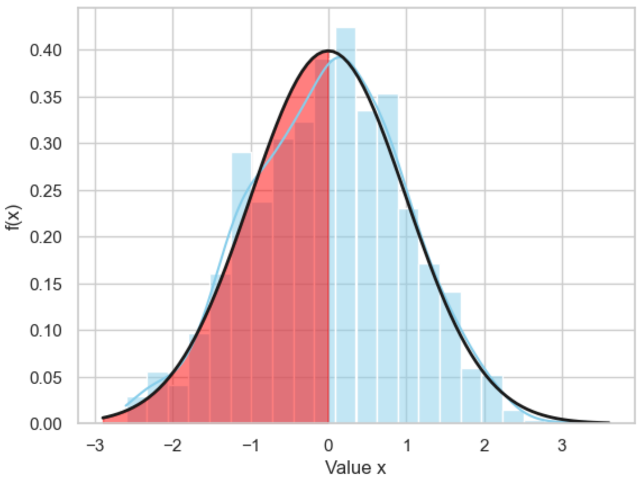
μ = expected value i.e. mean = 1 , 2, -1

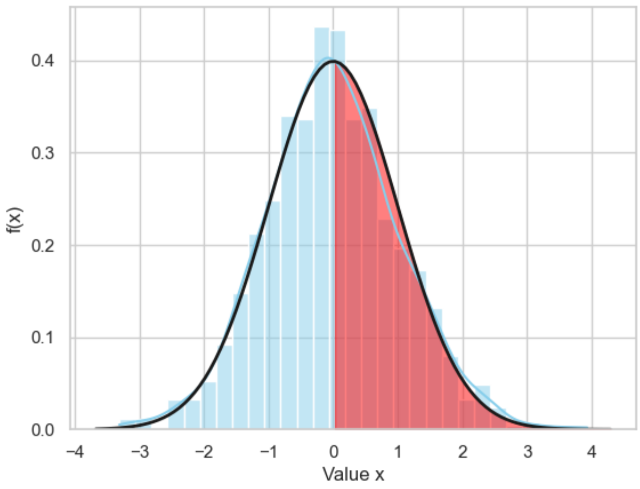
σ = Standard deviation = 1

Symmetric around Mean

- Pdf of normal distribution is given by $f(x) = \frac{1}{\sqrt{2\Pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$
- The probability density function of a normal random variable X exhibits symmetry around its mean.
- This means that X is equally likely to occur on either side of the mean (μ).

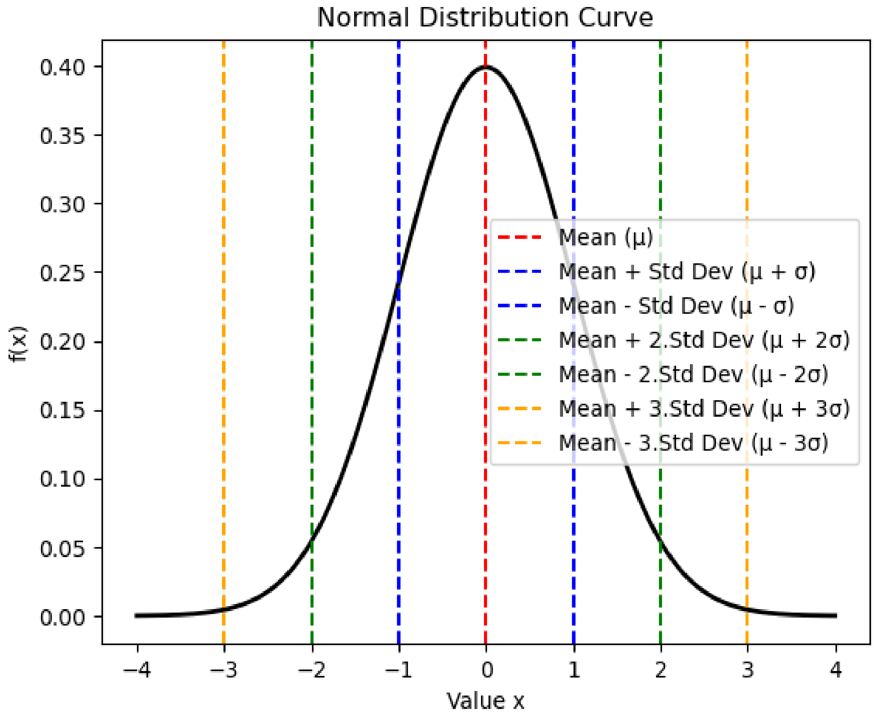
$$P(X \leq \mu) = P(X \geq \mu) = 0.5$$





Normal Distribution

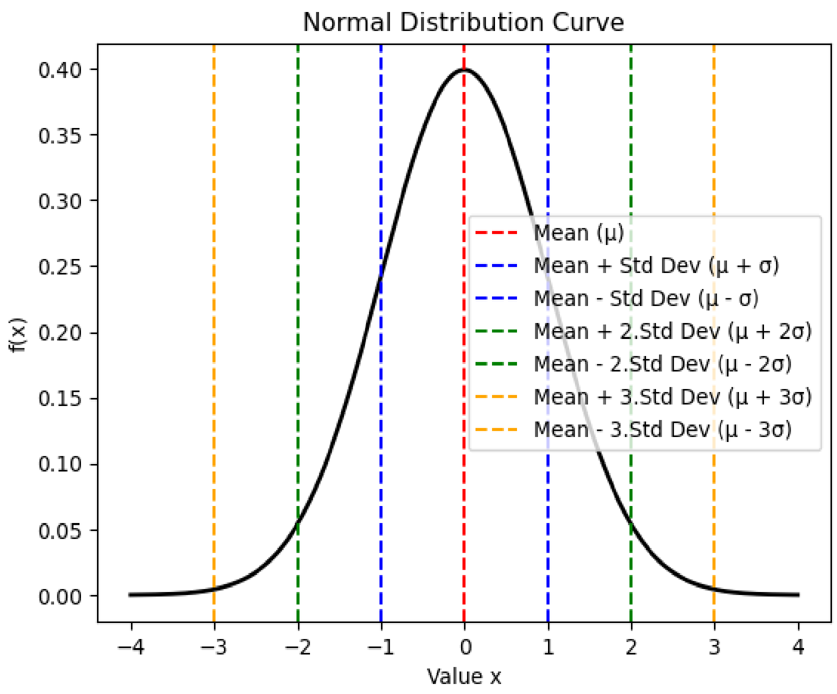
1. **Bell Shaped**
2. **Centered at mean i.e., the expected value**
3. **Close to the horizontal axis outside the range from μ + 3.σ and μ - 3.σ**



Approximation rule for Normal Distribution

For a continuous random variable following a normal distribution with mean μ and standard deviation σ:

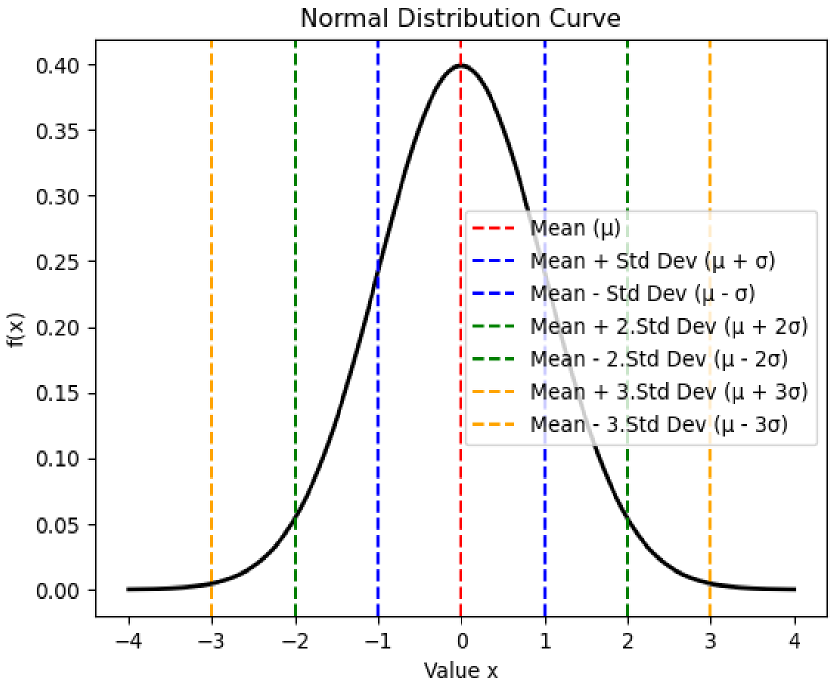
- **The interval between μ + σ and μ - σ encompasses approximately 68% of the distribution.**
- **The interval between μ + 2σ and μ - 2σ encompasses approximately 95% of the distribution.**
- **The interval between μ + 3σ and μ - 3σ encompasses approximately 99% of the distribution.**



Example of Normal Distribution N (μ , σ^2)
X~ N(10,16)

What is the probability that X is between 6 and 14 ?

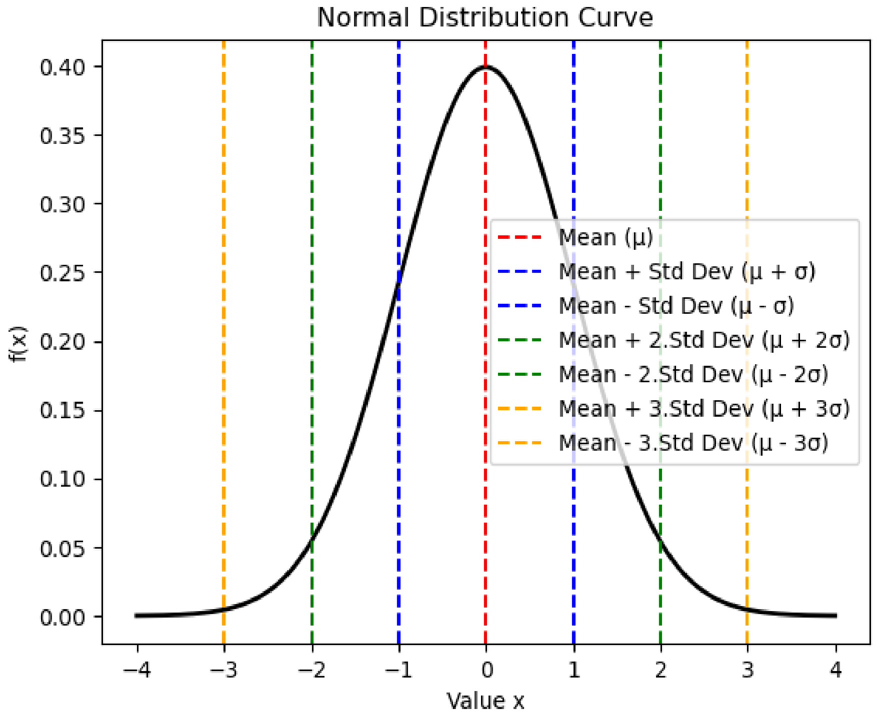
Using the Approximation rule Answer is 0.68



Example of Normal Distribution N (μ , σ^2)
X~ N(10,16)

What is the probability that X is less than 6 ?

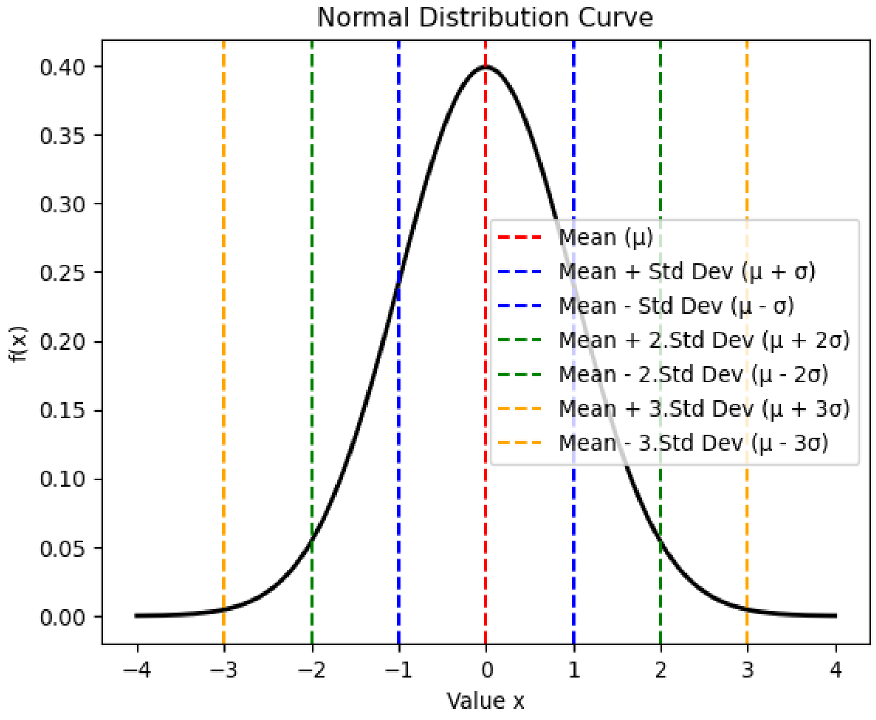
Using the Approximation rule Answer is 0.16



Example of Normal Distribution N (μ , σ^2)
X~ N(10,16)

What is the probability that X is greater than 14 ?

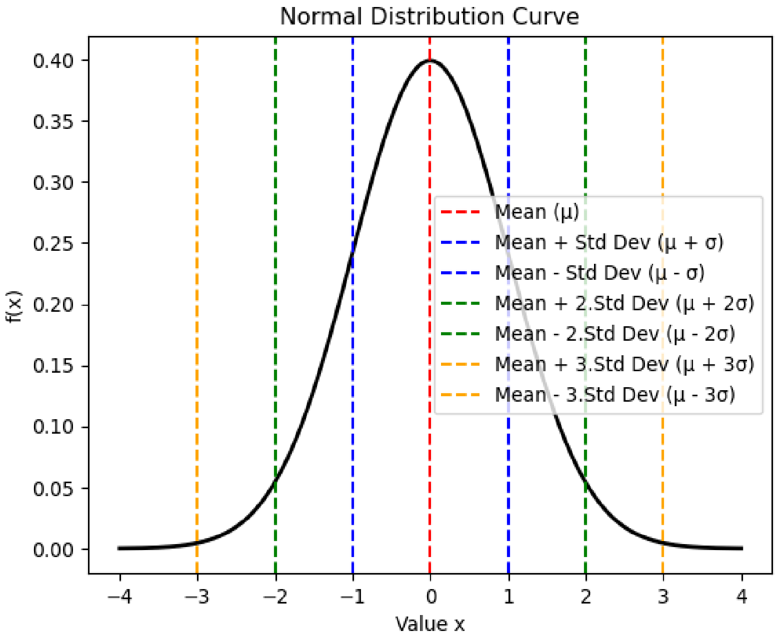
Using the Approximation rule, the Answer is 0.16



Example of Normal Distribution N (μ, σ^2)
X~ N(10,16)

What is the probability that X is greater than 22 ?

Using the Approximation rule, the Answer is 0



REFERENCES

- Introduction to Probability and Statistics for Engineers and Scientists, Sixth Edition, Sheldon M. Ross
- Statistical Methods Combined Edition (Volume I& II), N G Das

CONCLUSION

- Continuous random variable
- Probability Distribution Function
  Properties of pdf
- Cumulative Distribution Function
- Uniform and Standard Uniform distribution
  Expectation and Variance
- Exponential distribution
  Expectation and Variance
- Normal distribution
  Approximation Rule


THANK YOU

Mine Automation and Data Analytics