Continuous random variables

Outline

- Discrete vs continuous random variables
- Probability mass function vs Probability density function
- Properties of the pdf
- Cumulative distribution function
- Properties of the cdf
- Expectation, variance and properties
- The normal distribution

Recap

Till now, we discussed

Discrete random variables: can take a finite, or at most countably infinite, number of values,

For example:

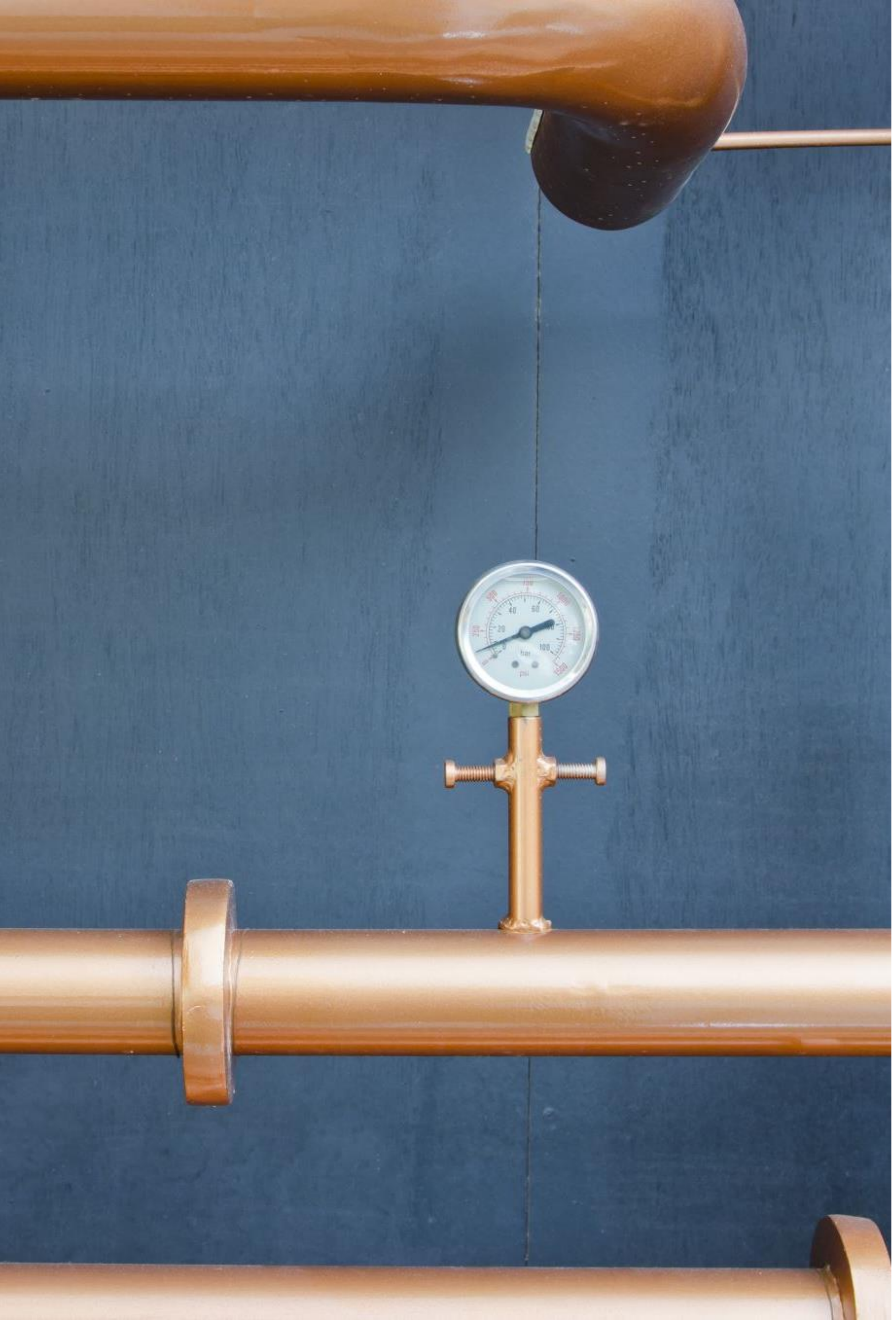
- Binomial random variable
- Bernoulli random variable
- Geometric random variable

Continuous random variable

- A continuous random variable is a random variable that Can take on an uncountably infinite range of values.

- Due to the above definition, the probability that a continuous random variable will take on an exact value is 0.

- For any specific value $X = x$, $P(X = x) = 0$



Continuous random variable

<u>Examples:</u>

- The volume of water passing through a pipe over a given time period.
- The height of a randomly selected individual.

Continuous random variable

$\int x^n dx = \frac{x^{n+1}}{n+1}$ $\int_{1/2}^{1} 4x^3 dx$

$= \left[\frac{4x^4}{4}\right]_{1/2}^{1}$

Example:

Suppose the **probability density function** of a continuous random variable, X, is given by $4x^3$, where $x \in [0, 1]$.

The probability that X takes on a value between 1/2 and 1 needs to be determined.

$\left(x^4\right)_{1/2}^{1}$

$1 - \frac{1}{16} = 15/16$

Solution

Solution:

This can be done by integrating $4x^3$ between 1/2 and 1. Thus, the required probability is

$$P = \int_{1/2}^{1} 4x^3$$

P = 15/16.

Probability density function (pdf)

- The probability density function (pdf) and the cumulative distribution function (CDF) are used to describe the probabilities associated with a continuous random variable.

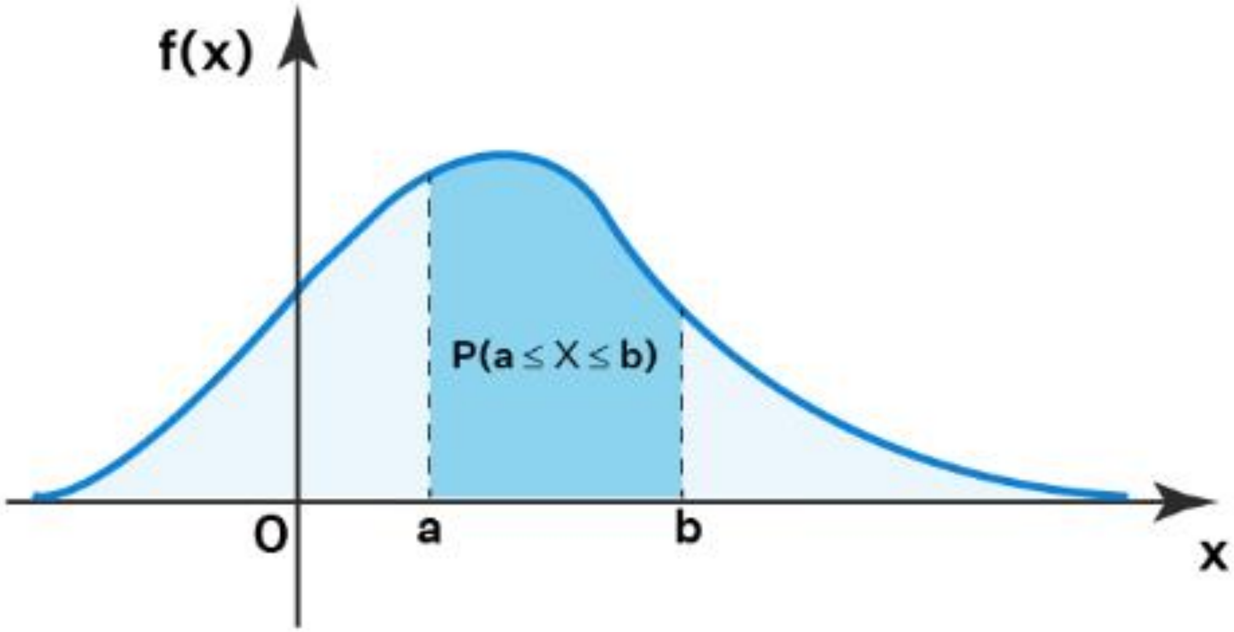
- For a continuous random variable, we cannot construct a PMF (discussed earlier for discrete) – each specific value has zero probability.

- Instead, we use a continuous, non-negative function $f_X\ (x)$ called the probability density function, or PDF, of X

Probability density function (pdf)

The probability of X lying between two values x_1 and x_2 is simply the area under the PDF, i.e

$$P(x_1 \le X \le x_2) = \int_{x_1}^{x_2} f_X(x)dx$$



Example

The pdf of a continuous random variable, X, is given as follows:

$$f(x) = \begin{cases} x & 0 \le x \le 1 \\ x + 3 & 1 < x \le 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that the value of the continuous random variable will lie between 0 and 0.5, i,e., Find $P(0 \le X \le 0.5)$.

Ans: 0.125

Solution

-

$$f(x) = \int_0^{0.5} f(x)\,dx$$

$$= \int_0^{0.5} x\,dx = \left.\frac{x^2}{2}\right|_0^{0.5}$$

$$= \frac{1}{2}\left(0.5 \times 5 - 0\right)$$

$$= 1/8 = 0.125$$

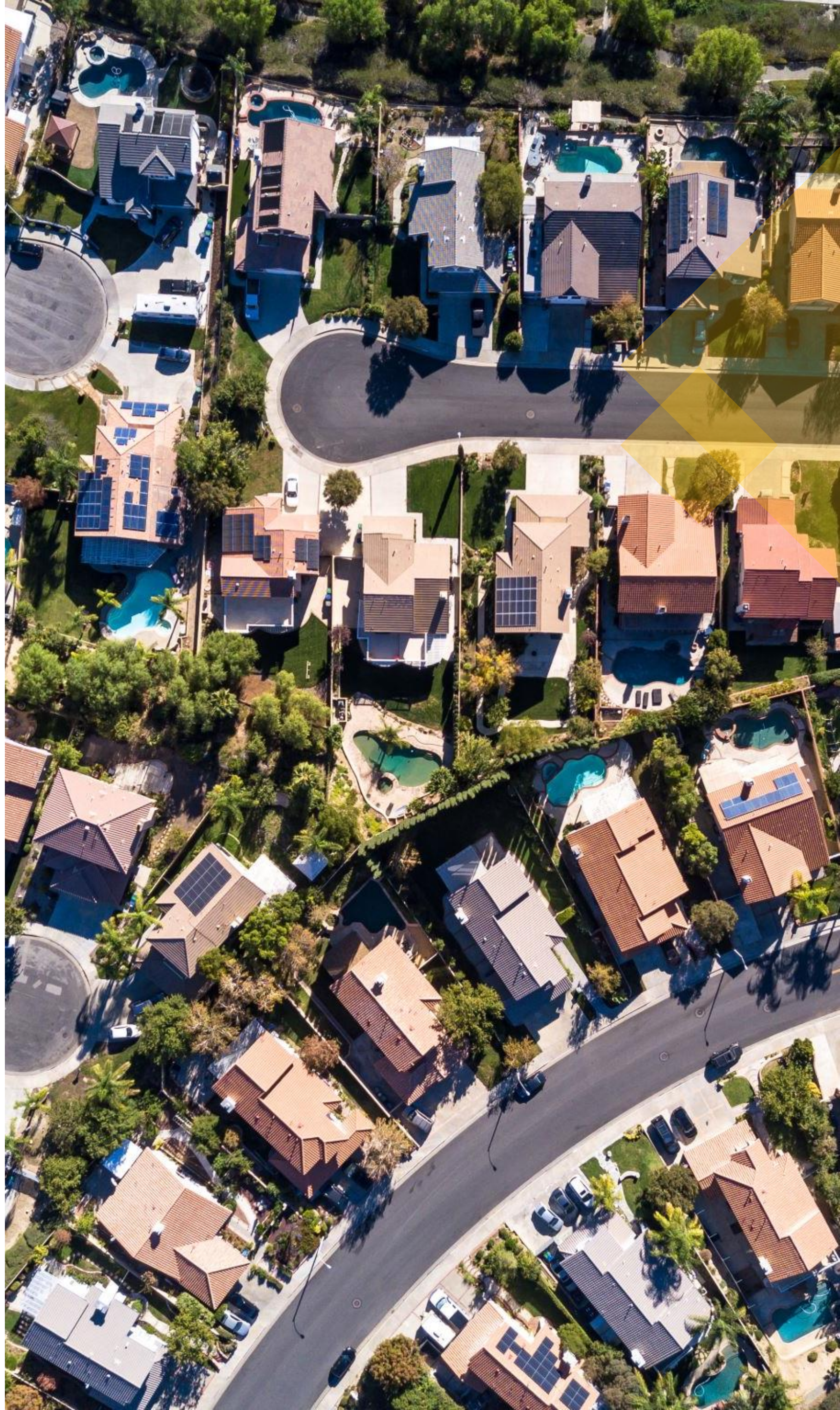
Properties of the pdf

- For any single value a, $P(X = a) = \int_{a}^{a} f_X(x)dx = 0$

- f(x) $\geq$ 0. This implies that the probability density function of a continuous random variable cannot be negative.

- $\int_{-\infty}^{-\infty} f_X(x)dx$ =1, this means that the total area under the graph of the pdf must be equal to 1.

- Note that fX (x) can be greater than 1 – even infinite! – for certain values of x, provided the integral over all x is 1.



Cumulative distribution function (cdf)

Now, we are interested in $P(X \leq x)$

Examples:

- What is the probability that the bus arrives before 1:30?
- What is the probability that a randomly selected person is under 5'7"

We can get this from our PDF:

$$F_X(x) = P(X \leq x') = \int_{-\infty}^{x'} f_X(x)dx$$

Note: If X is discrete, $f_X(x)$ is a piecewise-constant function of x

Cumulative distribution function (cdf)

- The CDF is monotonically non-decreasing:
if x ≤ y, then FX (x) ≤ FX (y)

- $F_X(x) \rightarrow 0$ as x → −∞

- $F_X(x) \rightarrow 1$ as x → ∞

Expectation of a continuous random variable

- Similar to the discrete case…

but we are integrating rather than summing

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

- Just as in the discrete case, we can think of E[X] as the "center of gravity" of the PDF

Expectation of a continuous random variable

Expectation of a function g(X) of a continuous random variable is

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

Note, g(X) can be a continuous random variable, e.g. $g(X) = X^2$, or a discrete random variable, e.g

$$g(X) = \begin{cases} 1 & \text{if } X \geq 0 \\ 0 & \text{if } X < 0 \end{cases}$$

Variance of a continuous random variable

$$var[X] = E[X^2] - E[X]^2$$

$$var[X] = \int_{-\infty}^{\infty} (x - E[x])^2 f_X(x) dx$$

Note:

$$E[aX + b] = aE[X] + b$$

$$var(aX + b) = a^2\, var(X)$$

Geometric distr: $E(x) = x p_i$

mean $\Rightarrow$ $1/p$

Variance $\Rightarrow$ q/p^2

S.D $\Rightarrow$ $\sqrt{q/p^2}$

$P(x) = q^{x-1} \cdot p$

Binomial Distr

mean $= p\,n$

var $= n\,p\,q$

S.D $= \sqrt{npq}$

${}^{n}C_x\, p^x\, q^{n-x}$

Poisson

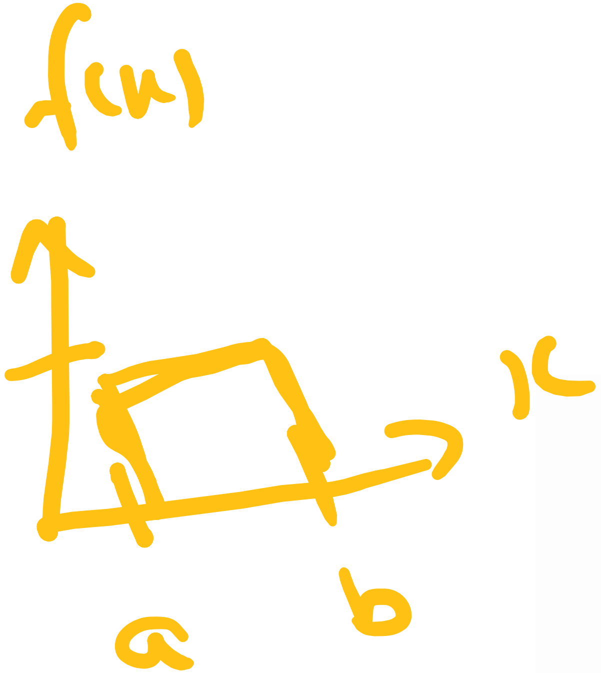
mean $= \frac{r}{p}$

var $= \frac{rq}{p}$

S.D $= \sqrt{\frac{rq}{p}}$

$\frac{\lambda^x e^{-\lambda}}{x!}$

Uniform Distribution

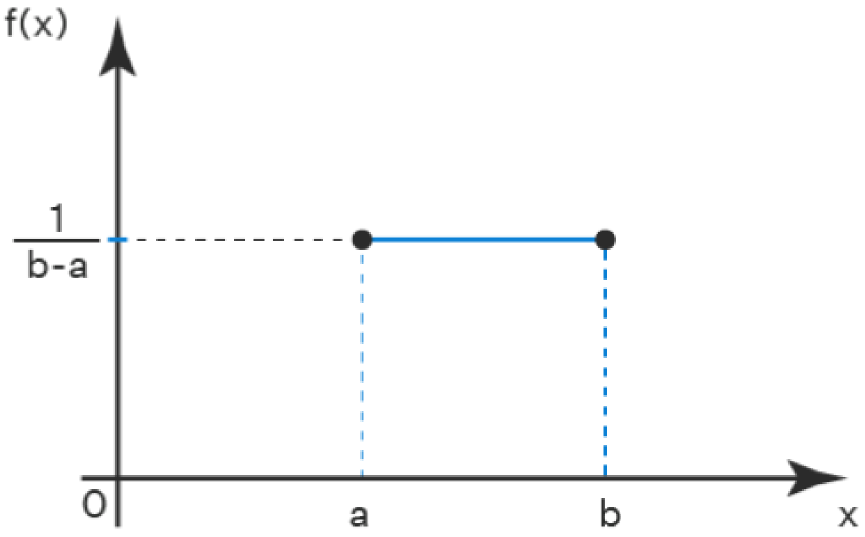


$$\Downarrow$$

Constant Probability within Domain

Uniform Random Variable

- A continuous random variable that is used to describe a uniform distribution is known as a uniform random variable.
- Such a distribution describes events that are equally likely to occur.
- The pdf of a uniform random variable is as follows:



$$f(x) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Mean of a uniform random variable

Let X be a uniform random variable over [a, b]. What is its expected value?

$$\mathrm{E}[X] = \int_{-\infty}^{\infty} x f_X(x) dx$$

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$E[X] = \int_{-\infty}^{a} x \times 0 dx + \int_{a}^{b} \frac{x}{b-a} dx + \int_{b}^{\infty} x \times 0 dx$$

$$\mathrm{E}[X] = \int_{a}^{b} \frac{x}{b-a} dx = \frac{a+b}{2}$$

Mean of a uniform random variable

Variance of a uniform random variable

$$var[X] = E[X^2] - E[X]^2$$

Variance of a uniform random variable

$$
\begin{aligned}
E[X^2] &= \int_{-\infty}^{\infty} x^2 f_X(x)dx \\
&= \int_a^b \frac{x^2}{b-a}dx \\
&= \left[\frac{x^3}{3(b-a)}\right]_a^b \\
&= \frac{b^3-a^3}{3(b-a)} = \frac{a^2+ab+b^2}{3}
\end{aligned}
$$

So, the variance is

$$
\mathrm{var}(X) = E[X^2] - E[X]^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}
$$

The normal distribution

A normal, or Gaussian, random variable is a continuous random variable with PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x-\mu)^2/2\sigma^2}$$

where,

μ = mean

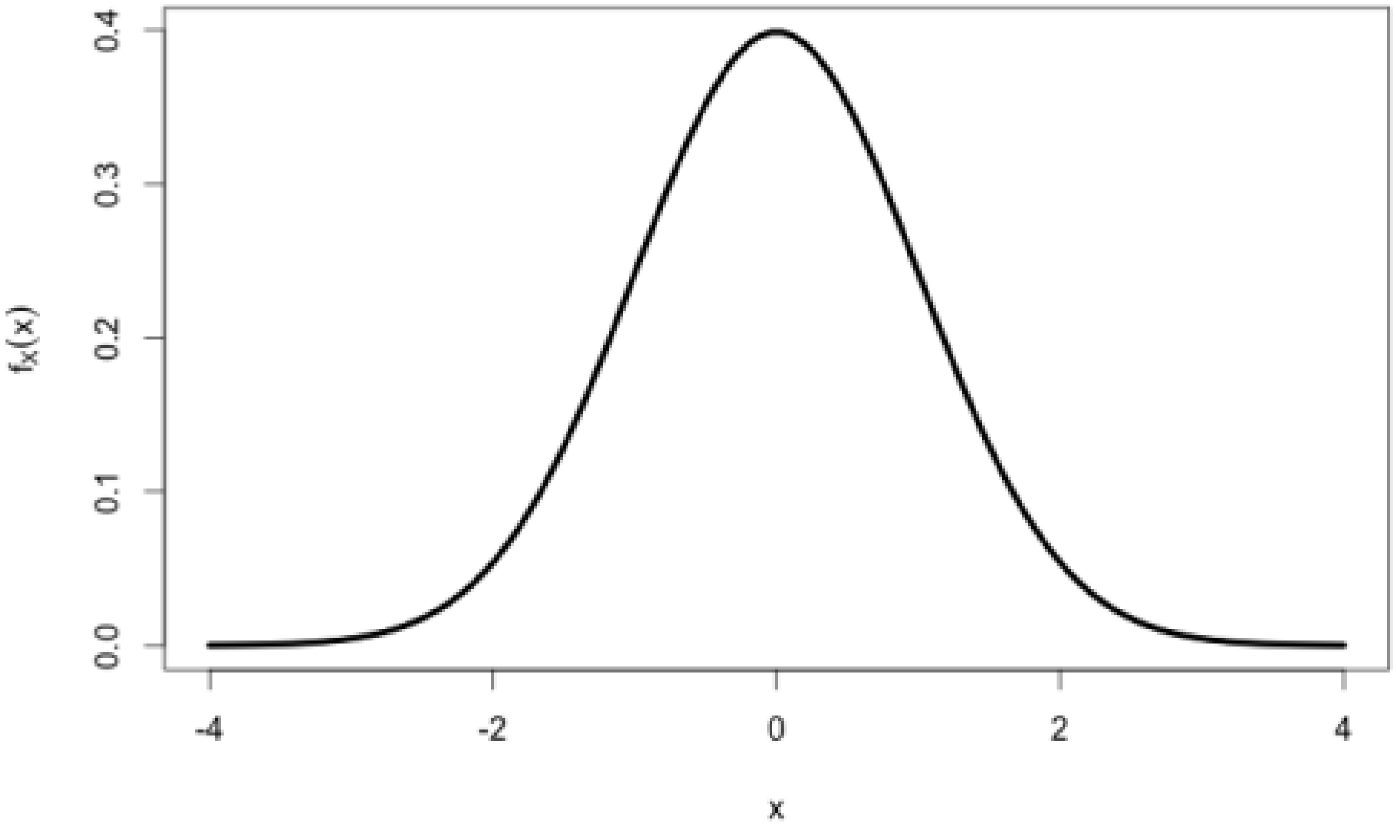
σ = standard deviation

$\sigma^2 =$ variance

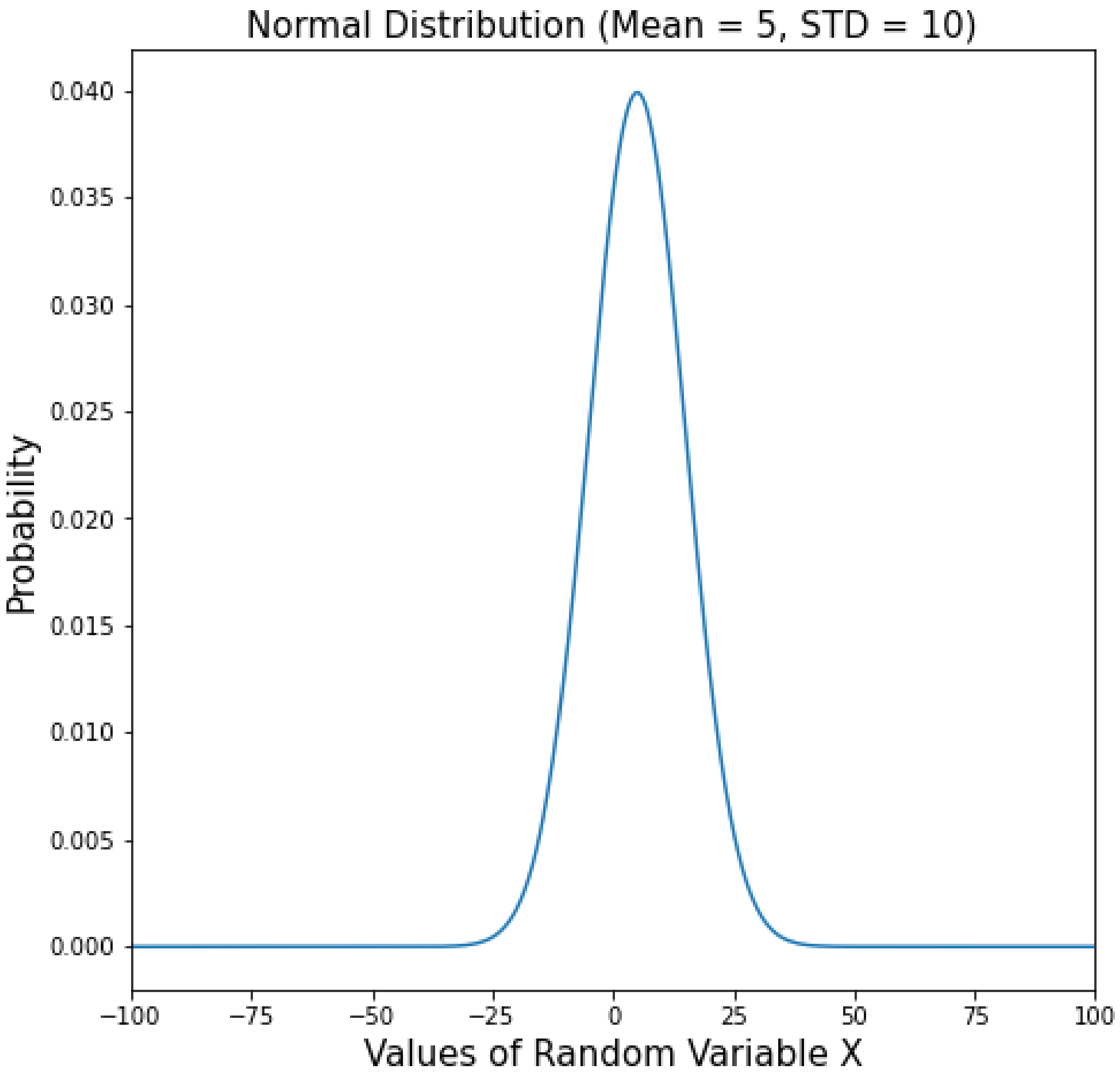
We write $X \sim N(\mu, \sigma^2)$

A normal distribution where $\mu = 0$ and $\sigma^2 = 1$ is known as a standard normal distribution

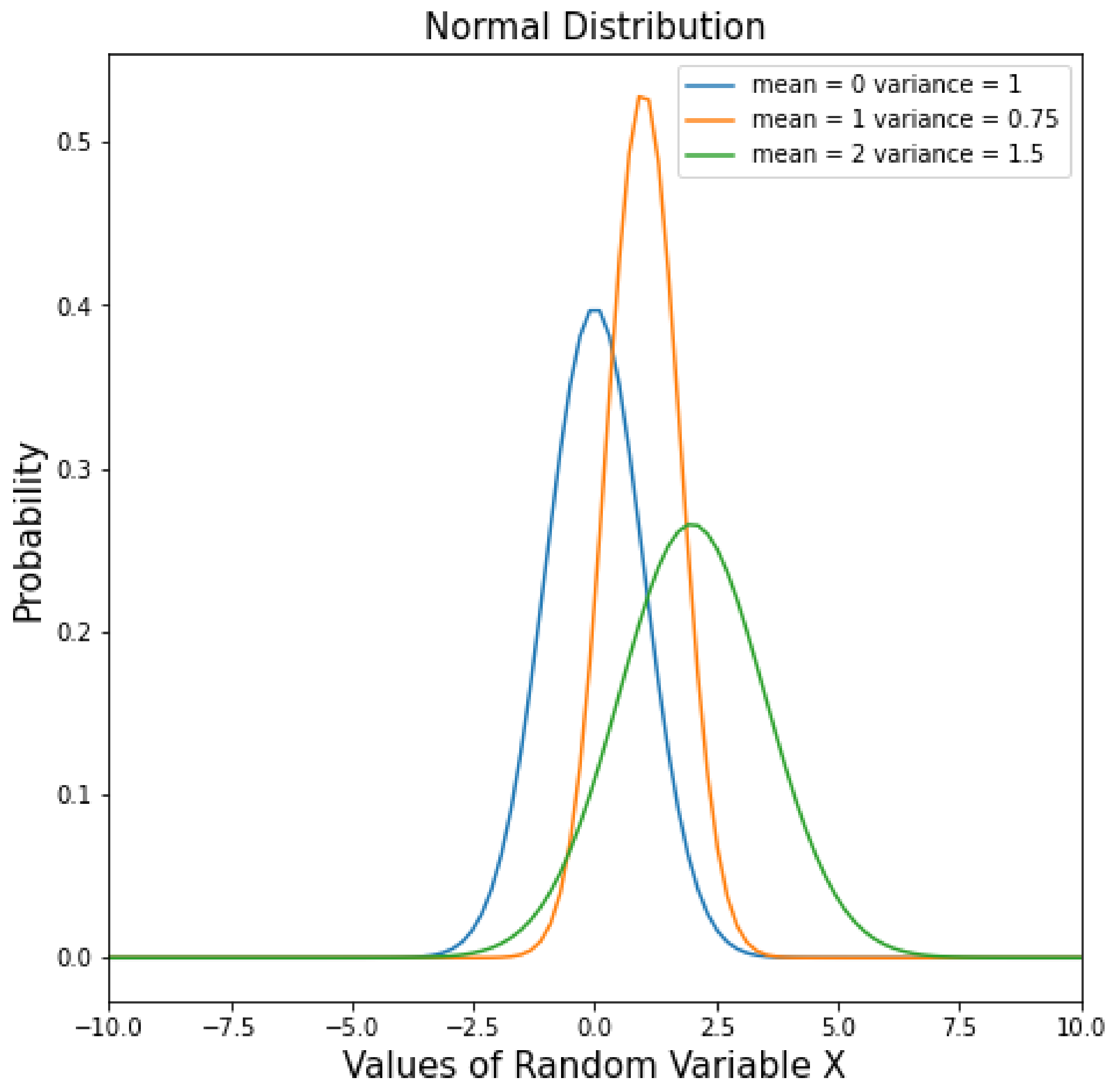
The normal distribution



The normal distribution



The normal distribution



The normal distribution

- The normal distribution is the classic "**bell-shaped curve**"

- Further, it has a number of nice properties that make it easy to work with. **Like symmetry**.

$P(X \geq 2) = P(X \leq -2)$ (from the curve)

Let $X \sim N(\mu, \sigma^2)$ and Let $Y = aX + b$

Then what are the mean and variance of Y ?

- $E[Y] = a\mu + b$
- $var[Y] = a^2\sigma^2$.
- Then Y is also a normal random variable with mean $a\mu + b$ and variance $var[Y] = a^2\sigma^2$.

The standard normal

- $X \sim N(0,1)$

- The pdf formula is as follows:

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

The standard normal

It is often helpful to map our normal distribution with mean and
variance 2 onto a normal distribution with mean 0 and variance 1.

If $X \sim N(\mu, \sigma^2)$, then $Z = \dfrac{X - \mu}{\sigma} \sim N(0, 1)$.

(Note, we often use the letter Z for standard normal random variables)

Mean and Variance of Normal Distribution

Example-1

If the height of 300 students are normally distributed with mean 64.5 inches and standard deviation 3.3 inches.

How many students have height

1. Less than 5 feet

2. Between 5 feet and 5 feet 9 inches

Example-2

The distribution of 500 workers in a factory is approximately Normal with mean and SD Rs 75 and Rs 15, respectively.

Find the No. of workers who receive weekly wages

1. More than 90

2. Less than 45