# UNIT - 2

**Distributions, Continuous Distributions, Probability, Decisions Trees, Machine Learning, Probability and Distributions, Descriptive Statistics and Graphs.**

R. K. Sharma, CSED, TIET Patiala

# Distributions, Continuous Distributions, Probability

# Basic Concepts

- Deterministic Vs Probabilistic Models
- Random Experiment: A random (or statistical) experiment is an experiment in which:

(a) All outcomes of the experiment are known in advance.

(b) Any performance of the experiment results in an outcome that is not known in advance.

(c) The experiment can be repeated under identical conditions.

# Basic Concepts:

- Sample Space:
  - The set of all possible outcomes of a random experiment is called the sample space (S) associated with that experiment.
- Finite and Infinite Sample Spaces:
  - If the set of all possible outcomes of the experiment is finite, then the associated sample space shall be called as finite sample space.
  - Otherwise, the sample space is called infinite sample space. Infinite sample spaces are further classified as countably infinite and uncountable spaces.
  - Finite sample space and countably infinite sample space together are called countable or discrete sample space.
  - If the sample space is uncountable, this is also called continuous sample space.

# Basic Concepts:

- Sample Space: Example
  - Throwing a coin, S = {H, T}
  - Throwing a die, S = {1, 2, 3, 4, 5, 6}
  - Finding the execution of a program, $S = \{t : t > 0\}$
  - Number of packets processed by a router, S = {0, 1, 2, …}
  - Marks of students in this course S = {0, 1, 2, 3, 4, …, 100}
  - …

# Basic Concepts:

- Event

  - *An event E is a subset of the sample space S.*
  - We know that if a set has n elements than it will have $2^n$ subsets.
  - *As such a sample space with n elements has* $2^n$ *events associated with it.*
  - Event {} is called impossible event and
  - the event S is called sure event.

# Basic Concepts:

- Mutually exclusive events
  - Two events $A$ and $B$ are said to be mutually exclusive if both these events cannot occur simultaneously.
  - Mathematically, it means that $A \cap B = \varphi$.
  - Mutually exclusive events are also called as disjoint events.

- Equally likely events
  - Two events $A$ and $B$ are said to be equally likely events if one has no reason to believe that one will occur more often than the other.

- Exhaustive set of events
  - A set of events $\{A_1, A_2, \ldots, A_k\}$ is called exhaustive if at least one of the events from the set must occur. Thus $\{A_1 \cup A_2 \cup \ldots, \cup A_k\} = S$ (Sample space)

# Basic Concepts:

- Statistical definition of probability
  - Suppose that an event $E$ occurs $m(E)$ times out of $n$ repetitions of a random experiment $R$. The ratio $m(E) / n$ is called the relative frequency of $E$, i.e., $r(E) = m(E) / n$
  - This has been observed with all random experiments that this relative frequency tends to a limit, i.e., as $n$ goes to infinity, this number tends to a specific number. This number is called the probability of the event $E$. Thus,

$$\lim_{n\to\infty} \frac{m(E)}{n} = P(E).$$

  - This definition of probability is called the statistical definition of probability.

# Basic Concepts:

- Mathematical definition of probability
  - If an experiment can result into $n$ mutually exclusive, equally likely and exhaustive ways, and out of which $m$ are favorable to an event $A$.
  - The probability of $A$, $P(A)$ is defined as:

    $P(A) = m / n$

# Basic Concepts:

- Axiomatic definition of probability
  - *A* real-valued function $P(A)$ defined on a finite sample space $S$ is a probability function if it satisfies the following axioms:
  - Axiom 1: $P(A) \geq 0$ for every event $A$ in $S$.
  - Axiom 2: $P(S) = 1$.
  - Axiom 3: $P(A \cup B) = P(A) + P(B)$ for every pair of events for which $A \cap B = \varphi$.
- As such, probability is a measure defined on the sample space. This is worth mentioning here that Statistical and Mathematical definitions follow these axioms.
- Axiom 3 above is stated as below in a situation when there are more than two events.
- Axiom 3(a): If $A_1, A_2, \ldots, A_i, \ldots$ is a countable sequence of pair-wise mutually exclusive events of a sample space $S$, then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$$

# Independent events

- Two events $A$ and $B$ are said to be independent if $P(A \cap B) = P(A).P(B)$.
- We can show that if $A$ and $B$ are independent events then, $A$ and $\bar{B}$; $\bar{A}$ and $B$ and $\bar{A}$ and $\bar{B}$ are also independent events.
  - Given that $P(A \cap B) = P(A).P(B)$, to prove: $P(A \cap \bar{B}) = P(A).P(\bar{B})$
  - Proof:
  - $P(A \cap B) = P(A).P(B) = P(A)\ (1- P(\bar{B})) = P(A) - P(A).P(\bar{B})$
  - So, $P(A) - P(A).P(\bar{B}) = P(A \cap B)$ …(a)
  - Also, $A = (A \cap \bar{B}) \cup (A \cap B)$ gives $P(A) = P(A \cap \bar{B}) + P(A \cap B)$ …(b)
  - Using (a) and (b), we have $P(A) - P(A).P(\bar{B}) = P(A) - P(A \cap \bar{B})$;
  - This gives $P(A \cap \bar{B}) = P(A).P(\bar{B})$

# Example

- Probability of guessing a correct password:
  - We have a set of $n$ passwords for your email. There are 2 situations: (i) password is selected without replacement, and (ii) password is selected with replacement, from the set of $n$ passwords. Obviously, one has to keep on guessing till a correct password is found and keep on counting the number of guesses. Describe the sample space for the two situations and the corresponding probabilities.

# Probability and Probability Distributions

- RANDOM VARIABLE (*X*)
  - A random variable (RV) *X* is a real valued function defined on the sample space S, i.e., for each element of the sample space, the RV takes a unique real number.

# Probability and Probability Distributions

- With a random variable $X$, we can associate a probability distribution.
- We need to define the probability distributions of discrete and continuous random variables separately.

- *Discrete Random Variable*
  - A set of numbers $\{p_n\}$ is the probability distribution of a random variable $X$ taking values $x_1, x_2, \ldots, x_n, \ldots$ if it satisfies the following properties:
  - $p_i \geq 0,\ i = 1, 2, \ldots$ (i)
  - $\sum_{i=1}^{\infty} p_i = 1$ (ii)
  - We also write this as $p(x) = p_i$, for $x = x_i$; 0, otherwise. This function $p(x)$ is also called the probability mass function (*pmf*) of the random variable X. This notation also means that,

    $P(X = x_i) = p_i,\ i = 1, 2, \ldots$
- We also say that a function that satisfies (i) and (ii) above is a legitimate *pmf*.

# Probability and Probability Distributions

- Continuous Random Variable
  - When we deal with a one-dimensional continuous random variable, we do not have the values of the variable at discrete points rather the random variable takes the values from an interval. This interval can be considered as a subset of the real line. Here, instead of probability mass function, we define probability density function. Probability of an event depends upon the starting point and the length of the interval. Probability density function (*pdf*) is defined as the function that gives us the probability per unit interval. This can be illustrated with the following.

$$f(x) = \lim_{h \to 0} \frac{P(x < X < x + h)}{h}$$

  - As such, for a continuous random variable *X*, we define a *pdf*. A function *f(x)* is said to be a *pdf* if it satisfies the following properties:
  - $f(x) \geq 0, -\infty < x < \infty$ (Support of Variable) ...(i)
  - $\int_{-\infty}^{\infty} f(x)dx = 1$ ...(ii)
  - This simply means that *pdf* always takes nonnegative values and the integral of this function over the interval under consideration for the random variable should be 1. We also say that a function that satisfies (i) and (ii) above is a legitimate *pdf*.

# Probability and Probability Distributions

- **Notes:**
  - Here, $P(a \le X \le b)$ is defined as $\int_a^b f(x)dx$.
  - For a continuous random variable, the probability that the variable takes a specific value is 0.
  - P(X = a ) = P(a <= X <= a ) = 0
- **Example:** Let us consider that *f(x)* = 2x, 0 < x < 1; 0, elsewhere. Let us show that this is a legitimate *pdf*.
  - We can see that $f(x) \ge 0$ for all $x$ such that $0 < x < 1$.
  - $\int_{-\infty}^{\infty} 2x\, dx = \int_0^1 2x\, dx = [x^2]_0^1 = 1$.
  - As such, this function is a legitimate *pdf*.

# Probability and Probability Distributions

- **Cumulative Distribution Function (or, Distribution Function)**
  - **The cumulative distribution function (*cdf*) of a random variable *X* is defined as,**
  - $F(t) = P(X \leq t), \; -\infty < t < \infty$
- For a discrete random variable, we can see that this function will be of the following form,
  - $F(t) = \sum_{x \leq t} p(x), \; -\infty < t < \infty$
- And for a continuous random variable, this function shall be,
  - $F(t) = P(X \leq t) = \int_{-\infty}^{t} f(x)dx$
- **Domain of F: Set of Real numbers,**
- **Range of F: $R^+$, 0 <= F(t) <= 1, F(-infinity) = 0, F(infinity) = 1**

# Probability and Probability Distributions

- Example on CDF:

# CHARACTERISTICS OF PROBABILITY DISTRIBUTIONS

- **Expectation**
  - **Definition 1:** Let $X$ be a random variable taking values $x_1, x_2, \ldots, x_n, \ldots$ with probabilities $p_1, p_2, \ldots, p_n, \ldots$, respectively, i.e., $P(X = x_i) = p_i$, $i = 1, 2, \ldots, n, \ldots$ . Then the mathematical expectation (or expectation) of $X$ is defined as,

    $$E(X) = \sum_{i=1}^{\infty} x_i p_i,$$

    provided that the series $\sum_{i=1}^{\infty} x_i p_i$ converges absolutely, i.e., $\sum_{i=1}^{\infty} |x_i| p_i$ is finite.

  - **Definition 2:** Let $X$ be a random variable with *pdf* $f(x)$, then $E(X)$ is defined as,

    $$E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx, \text{ if } \int_{-\infty}^{\infty} |x|\, f(x)\, dx \text{ is finite}$$

- Examples:

# CHARACTERISTICS OF PROBABILITY DISTRIBUTIONS

- Properties of Expectation

  If $X$ is a random variable and $g(X)$ is a function of $X$, then

  - $E[g(X)] = \sum g(x_i)p_i$, if $X$ is a discrete random variable and

    $E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx$, if $X$ is a continuous random variable.
  - $E(c) = c$, where $c$ is a constant.
  - $E(cX) = cE(X)$.
  - $E(a + bX) = a + bE(X)$, $a$ and $b$ are two constants.
  - $E(X + Y) = E(X) + E(Y)$, $X$ and $Y$ being two random variables.
  - $E(XY) = E(X).E(Y)$, if $X$ and $Y$ are two independent random variables.

# Moments of a Random variable

- Definition:
  - The $r^{th}$ moment of a random variable $X$ about a point $a$ is denoted by $\mu_r'$ and is defined as:

    $$\mu_r' = E(X-a)^r$$

  - If we take $a = 0$, this is the moment about origin and it will thus be:

    $$\mu_r' = E(X^r)$$

  - There are some moments that are of special interest. These are defined by considering point $a$ as the expected value of $X$ and these moments are called the central moments. We define $\mu_r$ as:

    $\mu_r = E(X-\mu)^r$ where $\mu = E(X)$.

  - When $r = 0$, we have $\mu_0 = 1$;
  - And for $r = 1$, we have $\mu_1 = 0$.
  - When $r = 2$, we have $\mu_2 = E(X-\mu)^2$.

# Moments of a Random variable

- We define $\mu_r$ as:

  $\mu_r = E(X-\mu)^r$ where $\mu$ = E(X).

  - When $r$ = 0, we have $\mu_0$ = 1;
  - And for $r$ = 1, we have $\mu_1$ = 0.
  - When $r$ = 2, we have $\mu_2 = E(X-\mu)^2$.
  - When $r$ = 3, we have $\mu_3 = E(X-\mu)^3$. This moment is the measure of asymmetry and is called as skewness.
  - Coefficient of Skewness = $\alpha_3 = \frac{\mu_3}{(\mu_2)^{3/2}}$
  - When $r$ = 4, we have $\mu_4 = E(X-\mu)^4$. This moment is the measure of kurtosis, i.e., whether the values are concentrated around the mean or are scattered. This measure defines the peakedness of the curve.
  - Coefficient of Kurtosis = $\alpha_4 = \frac{\mu_4}{(\mu_2)^2}$

# Median and Mode of a Distribution

- MEDIAN
  - Median means a central value.
  - Sample median *m* is a number that is exceeded by at most a half of observations and is preceded by at most a half of observations.
  - Population median M is a number that is exceeded with probability no greater than 0.5 and is preceded with probability no greater than 0.5. That is, median M is such that:

    $P(X \geq M) \leq 0.5$ and $P(X \leq M) \leq 0.5$

# Median and Mode of a Distribution

- **Discrete Distribution:**
  - Find $x_1$ and $x_2$ such that
  - $x_1$: the largest value of $X$ such that $P(X \le x_1) \le 0.5$
  - $x_2$: the smallest value of $X$ such that $P(X \le x_2) \ge 0.5$
  - Median (M) is $(x_1 + x_1) / 2$
- Example: Suppose the distribution is:

| X | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| P(X = x) | 0.2 | 0.1 | 0.1 | 0.3 | 0.3 |
| F(t) = P(X ≤ t) | 0.2 | 0.3 | 0.4 | 0.7 | 1.0 |

  - This gives $x_1 = 3$; and $x_2 = 4$. As such, $M = 3.5$

# Median and Mode of a Distribution

- Continuous Distribution:
  - Median (M) is such that
  - F(M) = 0.5 where F is the Cumulative Distribution Function associated with the *pdf* of the distribution.
- Example:
  - Consider the *pdf*, *f(x)* = 2*x*, 0 < *x* < 1; 0, elsewhere. CDF for this random variable is
  - $F(t) = \ 0,\ t < 0$
  - $t^2,\ 0 \leq t \leq 1$
  - $1,\ t > 1$
  - This gives us, M = 0.7071

# Median and Mode of a Distribution

- MODE
  - The mode or modal value of a discrete random variable $X$ is the value $x$ for which the probability mass function (*pfm*) takes its maximum value.
  - The mode or modal value of a continuous random variable $X$ with a probability density function $f(x)$ is the value of $x$ for which $f(x)$ takes a maximum value. Thus mode is the $x$-coordinate of the maximum point on the graph of $f(x)$.

# Special Distributions: Discrete and Continuous

# Binomial Distribution

- Let us first understand what is a Bernoulli trial.
- A sequence of trials is called a Bernoulli trial if this follows the following characteristics.
  - Each trial results into one of the only two outcomes. These outcomes are generally called 'success' and 'failure'.
  - Probability of success for each trial is same. Let us denote this by *p*.
  - The sequence of trials contains finite number of elements. Let us denote this by *n*.
  - The trials are independent with each other.
- Let us now understand the Binomial distribution. Suppose that a sequence of *n* Bernoulli trials is given. Let *X* be the number of successes in *n* trials. Then *X* is a random variable that can take the values 0, 1, 2, …, *n*. This variable is called the Binomial variable. *X* shall follow the distribution,

  $P(X=x) = {}^{n}c_{x}p^{x}q^{n-x}$, where $q = 1 - p$, $x = 0, 1, 2, \ldots, n$.

- One may note that *p* is the probability of success and *q* is the probability of failure here. The random variable *X* is said to have Binomial distribution. This distribution, as one can see, needs two constants to define it completely. These are *n* and *p*. These constants are called the parameters of the distribution.
- As such, Binomial distribution is a two parameter distribution.

# Binomial Distribution

- A discrete random variable $X$ is said to follow Binomial distribution if its *pmf* is of the form,

$$P(X = x) = p(x) = {}^{n}c_{x} p^{x} q^{n-x}, \; x = 0, 1, 2, \ldots, n$$

where $q = 1 - p$, $p > 0$

# Binomial Distribution

### *Expected value of Binomial distribution*

- If $X$ follows Binomial distribution with parameters $p$ and $n$, its expected value (mu) shall be:
- $E(X) = \sum_{x=0}^{n} x \, {}^{n}c_x p^x q^{n-x}$
- $= \sum_{x=1}^{n} x \, \frac{(n)!}{(x)!(n-x)!} \, p^x q^{n-x}$
- $= \sum_{x=1}^{n} \frac{(n)!}{(x-1)!(n-x)!} \, p^x q^{n-x}$
- $= \sum_{x=1}^{n} \frac{n(n-1)!}{(x-1)!(n-x)!} p \, p^{x-1} q^{n-x}$
- $= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} \, p^{x-1} q^{n-x}$
- $= np \, (p+q)^{n-1}$
- $= np$

# Binomial Distribution

***Variance of Binomial distribution***

- Let us first show that for a random variable $X$ , $Var(X) = E(X^2) - (E(X))^2$.
- We have defined that $Var(X) = \sigma^2 = E\big(X - E(X)\big)^2$
- $= E(X^2 - 2XE(X) + (E(X))^2)$
- $= E(X^2) - 2E(X)E(X) + (E(X))^2$
- $= E(X^2) - (E(X))^2$
- Using the similar steps as illustrated in the expectation of a Binomial variable, we can obtain that for a Binomial random variable with parameters *n* and *p*,
- $\boldsymbol{E(X^2) = n(n-1)p^2 + np.}$
- $Var(X) = \sigma^2 = n(n-1)p^2 + np - n^2p^2$
- $= np - np^2$
- $= npq$

# Binomial Distribution

**Example:**

- Suppose the probability that a chip produced by a machine is defective is 0.2. If 10 such chips are selected at random, what is the probability that not more than one defective chip is found in these chips?

# *Poisson Distribution*

- ***Poisson Distribution***
- A random variable $X$ is said to follow Poisson distribution if its *pmf* is given by,
- $$p(x) = P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!},\ x = 0, 1, 2, \ldots;\ \lambda > 0.$$
- $Parameter\ is\ Lambda$

# Poisson Distribution

## Legitimacy of the Probability Mass Function

- One can note that $\frac{e^{-\lambda}\lambda^{x}}{x!}$ takes non-negative values for x = 0, 1, 2, … and λ > 0.
- Also,
- $\sum_{x=0}^{\infty}\frac{e^{-\lambda}\lambda^{x}}{x!}$
- $= e^{-\lambda}\sum_{x=0}^{\infty}\frac{\lambda^{x}}{x!}$
- $= e^{-\lambda}e^{\lambda}$
- = 1
- As such, $p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$, x = 0, 1, 2, …; λ > 0 is a legitimate probability mass function. It is worth noting that Poisson distribution involves one parameter.

# Poisson Distribution

***Expected value of Poisson distribution***

- If a random variable *X* follows Poisson distribution with parameter λ, then its expected value can be obtained as:
- $E(X) = \sum_{x=0}^{\infty} x\, e^{-\lambda}\lambda^{x}/(x)!$
- $= e^{-\lambda} \sum_{x=0}^{\infty} x\, \lambda^{x}/(x)!$
- $= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \lambda^{x-1}/(x-1)!$
- $= \lambda e^{-\lambda} \sum_{y=0}^{\infty} \lambda^{y}/(y)!$
- $= \lambda e^{-\lambda} e^{\lambda}$
- $= \lambda$
- As such, the expected value of the Poisson distribution is equal to the value of the parameter λ.

***Variance of Poisson distribution***

- If a random variable *X* follows Poisson distribution with parameter λ, then we can obtain the value of $E(X^2)$ using the similar idea as above. This will come out that $E(X^2) = \lambda^2 + \lambda$. As such, $Var(X) = E(X^2) - (E(X))^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$. As such the expected value and variance of Poisson distribution is same and is equal to the value of the parameter of the distribution.

# Poisson Distribution

***Relationship with Binomial and Poisson Distributions***

Poisson distribution can be obtained as a limiting form from Binomial distribution. In a Binomial distribution, if the number of trials *n* tends to infinity and the probability of success *p* tends to zero such that their product *np* tends to a finite quantity, say, λ, then the Binomial distribution tends to Poisson distribution with parameter λ.

- *Proof:*
- For a Binomial distribution, the *pmf* is given by
- $p(x) = {}^{n}c_x p^x q^{n-x}$, x = 0, 1, 2, …, *n*, where *q* = 1 – *p* (> 0)
- Let us impose the limits on *n* and *p* as above. In that limiting case, this *pmf* shall be:
- $p(x) = \lim_{n\to\infty, p\to 0\ (with\ np=\lambda)} \left(\frac{n(n-1)(n-2)\ldots(n-x+1)}{x!}\right) p^x q^{n-x}$
- $= \lim_{n\to\infty} \left(\frac{(n-0)(n-1)(n-2)\ldots(n-(x-1))}{x!}\right) \left(\frac{\lambda}{n}\right)^x (1-\frac{\lambda}{n})^{n-x}$
- $= \lim_{n\to\infty} \left(\frac{1(1-1/n)(1-2/n)\ldots(1-x/n+1/n)}{x!}\right) \lambda^x (1-\frac{\lambda}{n})^n / (1-\frac{\lambda}{n})^x$
- $= \frac{1}{x!}\lambda^x e^{-\lambda}$
- As such the *pmf* becomes *p*(x) = $\frac{e^{-\lambda}\lambda^x}{x!}, x = 0, 1, 2, \ldots$ in the limiting case. This is nothing but the *pmf* of the Poisson distribution.

# Poisson Distribution

**Example:**

Suppose the probability that a chip produced by a machine is defective is 0.2. If 10 such chips are selected at random, what is the probability that not more than one defective chip is found in these chips?

# Rectangular Distribution (or Uniform Distribution)

- A continuous random variable *X* is said to follow uniform distribution if its *pdf* is given by,
- $f(x) = \frac{1}{b-a}, a < x < b; 0, elsewhere$
- How to find the expected value and variance of Uniform Distribution.

**Example:** A point is randomly selected from the interval (4,10). What is the probability that this point shall lie in the interval (4.5, 6.5)?

# Exponential Distribution

- A continuous random variable $X$ is said to follow exponential distribution if its *pdf* is given by,
- $$f(x) = \lambda e^{-\lambda x}, x \geq 0, \quad \lambda > 0.$$
- Here, $\lambda$ is the parameter associated with exponential distribution.

# Exponential Distribution

***Cumulative Distribution Function of Exponential Distribution***

- Cumulative distributive function of exponential distribution $F(t)$ takes a very explicit form. This can be obtained using the definition.
- $F(t) = P(X \leq t)$
- $= \int_{-\infty}^{t} f(x)dx$
- $= \int_{o}^{t} \lambda e^{-\lambda x} dx$, using the *pdf* of exponential distribution
- $= [-e^{-\lambda x}]_0^t = [e^{-\lambda x}]_t^0 = 1 - e^{-\lambda t}$

- As such, the cumulative distribution function of exponential distribution is given by,
- $F(t) = 0, t \leq 0;\ 1 - e^{-\lambda t}, t > 0.$

Or (equivalently)

- $F(x) = 0,\ x \leq 0;\ 1 - e^{-\lambda x}, x > 0.$

# Exponential Distribution

***Expected value and Variance of Exponential Distribution: Do Yourself***

***Lack of Memory Property of Exponential Distribution***

- Exponential distribution possesses a very interesting property called as no memory property. This probability can mathematically be expressed as below.
- Suppose that $X$ is a random variable following exponential distribution. Then for $s, t > 0$, we have,
- $$P(X > s + t \mid X > s) = P(X > t).$$
- Few variables that may follow exponential distribution:
  - **Life of a light bulb, The age of human beings, execution time of a program, processing time of the packets, the time in loading of a website, …**

# Normal Distribution

- This distribution is an important distribution owing to the fact that a number of real life variables can be thought of following this distribution and the fact that in a limiting situation, the *pdf* of a number of distributions tend to the *pdf* of normal distribution. This distribution is a **two parameter** family of distributions.
- A continuous random variable *X* is said to follow normal distribution if its *pdf* is given by,
- $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$
- $$-\infty < \mu < \infty$$
- $$\sigma > 0$$
- $-\infty < x < \infty$ is the support of distribution.
- Symbolically, we write this as $X \sim N(x, \mu, \sigma^2)$.

# Normal Distribution

- This distribution is an important distribution owing to the fact that a number of real life variables can be thought of following this distribution and the fact that in a limiting situation, the *pdf* of a number of distributions tend to the *pdf* of normal distribution. This distribution is a **two parameter** family of distributions.
- A continuous random variable *X* is said to follow normal distribution if its *pdf* is given by,
- $$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty$$
- $$-\infty < \mu < \infty$$
- $$\sigma > 0$$
- $-\infty < x < \infty$ is the support of distribution.
- Symbolically, we write this as $X \sim N(x, \mu, \sigma^2)$.
  - **Legitimacy of this pdf: Do Yourself.**

# Normal Distribution

## Properties of Normal Distribution

- The *pdf* $f(x)$ of normal distribution takes maximum value at $x = \mu$. Prove this. It means that *Mode* of the distribution is mu.
- The *pdf* $f(x)$ of normal distribution is symmetrical about the line $x = \mu$.
- Points of inflexion of *pdf* of normal distribution are $\mu \pm \sigma$.
- The *pdf* of Normal distribution is a bell-shaped curve.
- All moments of odd order of normal distribution are zero.

# Normal Distribution

***Problem***

- Suppose that X is a random variable, and Y = **aX + b**. [Scaling and Shifting]
- Find E(Y) and Var(Y) in terms of E(X) and Var(Y).
  - **Solution:**
- **Expected Value and Variance of Normal Distribution: Do yourself.**
  - **If $X \sim N(x, \mu, \sigma^2)$, $E(X) = \mu$, and $V(X) = \sigma^2$ : Prove this.**
  - **If $X \sim N(x, \mu, \sigma^2)$, then $E\left( (X-\mu)/\sigma \right) = 0$; and $V\left( (X-\mu)/\sigma \right) = 1$**
  - **This process is called "Standardizing" a Variable.**
- Example: The marks of students $(X)$ in a class of 70 students follows normal distribution with mean 50 units and variance 225 units. Find the probability that $P(40 < X < 60)$.

Thank You !