UNIT-I

Probability and Random Variable

Introduction to Set

- Set: A set is a well defined collection of objects. These objects are called elements or members of the set. Usually uppercase letters are used to denote sets.
- The set theory was developed by George Cantor in 1845-1918. Today, it is used in almost every branch of mathematics and serves as a fundamental part of present-day mathematics.
- In everyday life, we often talk of the collection of objects such as a bunch of keys, flock of birds, pack of cards, etc.
- In mathematics, we come across collections like natural numbers, whole numbers, prime and composite numbers.

Laws in set theory

- $A \cap B = B \cap A$ (Commutative law)
- $(A \cap B) \cap C = A \cap (B \cap C)$ (Associative law)
- $\Phi \cap A = \Phi$ (Law of Φ)
- $U \cap A = A$ (Law of U)
- $A \cap A = A$ (Idempotent law)
- $A \cap (B U C) = (A \cap B) U (A \cap C)$ (Distributive law) Here $\cap$ distributes over U
- Also, $A U (B \cap C) = (A U B) \cap (A U C)$ (Distributive law) Here U distributes over $\cap$

Probability

- Experiment:

  In probability theory, an experiment or trial (see below) is any procedure that can be infinitely repeated and has a well-defined set of possible outcomes, known as the sample space.

- An experiment is said to be random if it has more than one possible outcome, and deterministic if it has only one.
- A random experiment that has exactly two (mutually exclusive) possible outcomes is known as a Bernoulli trial.

Experiment

Experiment	Outcomes
Flip a coin	Heads, Tails
Exam Marks	Numbers: 0, 1, 2, ..., 100
Assembly Time	$t > 0$ seconds
Course Grades	F, D, C, B, A, A+

Random Experiment

- An experiment is a random experiment if its outcome cannot be predicted precisely. One out of a number of outcomes is possible in a random experiment.
- A single performance of the random experiment is called a trial.Random experiments are often conducted repeatedly, so that the collective results may be subjected to statistical analysis.
- A fixed number of repetitions of the same experiment can be thought of as a composed experiment, in which case the individual repetitions are called trials.
- For example, if one were to toss the same coin one hundred times and record each result, each toss would be considered a trial within the experiment composed of all hundred tosses.

Relative frequency, Experiments

- Relative Frequency:

  Random experiment with sample space S. we shall assign non-negative number called probability to each event in the sample space. Let A be a particular event in S. then "the probability of event A" is denoted by P(A).

- Suppose that the random experiment is repeated n times, if the event A occurs n_A times, then the probability of event A is defined as "Relative frequency

- Event A is defined as

$$P(A) = \lim_{n \to \infty} \frac{n_A}{n}$$

Sample Space

- Sample Space: The sample space is the collection of all possible outcomes of a random experiment. The elements of are called sample points. A sample space may be finite, countable infinite or uncountable.
- A list of exhaustive *don't leave anything out] and mutually exclusive outcomes [impossible for 2 different events to occur in the same experiment] is called a sample space and is denoted by S.
- The outcomes are denoted by $O_1, O_2, \ldots, O_k$
- Using notation from set theory, we can represent the sample space and its outcomes as:

$$S = \{O_1, O_2, \ldots, O_k\}$$

Sample Space

- Given a sample space $S = \{O_1, O_2, \ldots, O_k\}$, the probabilities assigned to the outcome must satisfy these requirements:

(1) The probability of any outcome is between 0 and 1

i.e. $0 \leq P(O_i) \leq 1$ for each *i*, and

(2) The sum of the probabilities of all the outcomes equals 1

i.e. $P(O_1) + P(O_2) + \ldots + P(O_k) = 1$

Discrete and Continuous Sample Spaces

- Probability assignment in a discrete sample space: Consider a finite sample space . Then the sigma algebra is defined by the power set of S. For any elementary event   , we can assign a probability such that, For any event , we can define the probability

Continuous sample space

- Suppose the sample space *S* is continuous and uncountable. Such a sample space arises when the outcomes of an experiment are numbers. For example, such sample space occurs when the experiment consists in measuring the voltage, the current or the resistance.

Events

- The probability of an event is the sum of the probabilities of the simple events that constitute the event.
- E.g. (assuming a fair die) S = {1, 2, 3, 4, 5, 6} and P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6
- Then: P(EVEN) = P(2) + P(4) + P(6) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2

Types of Events

1. Exhaustive Events:

   A set of events is said to be exhaustive, if it includes all the possible events. Ex. In tossing a coin, the outcome can be either Head or Tail and there is no other possible outcome. So, the set of events{ H , T } is exhaustive.

2. Mutually Exclusive Events:

   Two events, A and B are said to be mutually exclusive if they cannot occur together. i.e. if the occurrence of one of the events precludes the occurrence of all others, then such a set of events is said to be mutually exclusive. If two events are mutually exclusive then the probability of either occurring is

Types of Events

3. Equally Likely Events:

If one of the events cannot be expected to happen in preference to another, then such events are said to be Equally Likely Events.(Or) Each outcome of the random experiment has an equal chance of occurring.

Ex. In tossing a coin, the coming of the head or the tail is equally likely

4. Independent Events:

Two events are said to be independent, if happening or failure of one does not affect the happening or failure of the other. Otherwise, the events are said to be dependent.

Probability Definitions and Axioms

Relative frequency Definition:

Consider that an experiment E is repeated n times, and let A and B be two events associated with E. Let n_A and n_B be the number of times that the event A and the event B occurred among the n repetitions respectively. The relative frequency of the event A in the 'n' repetitions of E is defined as

$$f(A) = nA / n$$

Axioms of Probability

- The Relative frequency has the following properties:
- $0 \leq f(A) \leq 1$
- $f(A) = 1$ if and only if A occurs every time among the n repetitions.
- If an experiment is repeated n times under similar conditions and the event A occurs in n_A times, then the probability of the event A is defined as

Joint probability

- Joint probability:

  Joint probability is defined as the probability of both A and B taking place, and is denoted by P (AB) or $P(A \cap B)$.

- probability notation: $P(AB) = P(A \mid B) * P(B)$

Conditional Probability

- Conditional probability is used to determine how two events are related; that is, we can determine the probability of one event given the occurrence of another related event.
- Experiment: random select one student in class.
- P(randomly selected student is male)
- P(randomly selected student is male/student is on 3^{rd} row)
- Conditional probabilities are written as P(A | B) and read as "the probability of A given B" and is calculated as

$$P(A \mid B) = \frac{P(A \text{ and } B)}{P(B)}$$

Bayes' Theorem

- Bayes' Law is named for Thomas Bayes, an eighteenth century mathematician.
- In its most basic form, if we know P(B | A),
- we can apply Bayes' Law to determine P(A | B)
- Bayes' theorem centers on relating different conditional probabilities. A conditional probability is an expression of how probable one event is given that some other event occurred (a fixed value).
- For a joint probability distribution over events A and B , P(A^B), the conditional probability of given is defined as

Bayes' theorem

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}.$$

- Note that P(A^B) is the probability of both A and B occurring, which is the same as the probability of A occurring times the probability that B occurs given that A occurred P(B/A)*P(A)
- Using the same reasoning P(A^B), is *also* the probability that B occurs times the probability that A occurs given that B occurs: P(A/B)*P(B) The fact that these two expressions are equal leads to Bayes' Theorem. Expressed mathematically, this is:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}, \text{ if } P(B) \neq 0,$$

$$P(B \mid A) = \frac{P(B \cap A)}{P(A)}, \text{ if } P(A) \neq 0,$$

$$\Rightarrow P(A \cap B) = P(A \mid B) \times P(B) = P(B \mid A) \times P(A),$$

$$\Rightarrow P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B)}, \text{ if } P(B) \neq 0.$$

Bayes' theorem

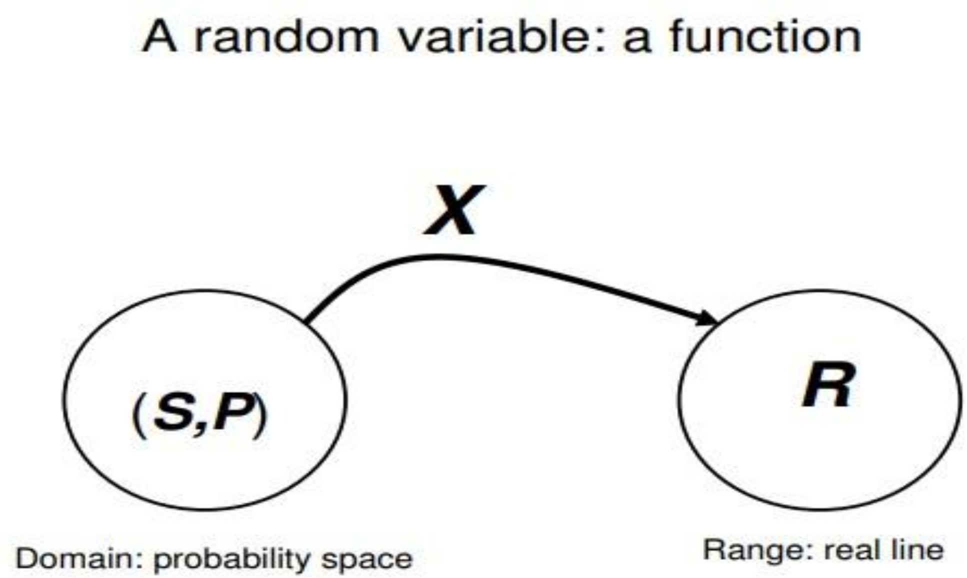
- The probabilities P(A) and P(A^C) are called prior probabilities because they are determined prior to the decision about taking the preparatory course.
- The conditional probability P(A | B) is called a posterior probability (or revised probability), because the prior probability is revised after the decision about taking the preparatory course.

Random variable

- A (real-valued) random variable, often denoted by X (or some other capital letter), is a function mapping a probability space (S, P) into the real line R. This is shown in next slide.

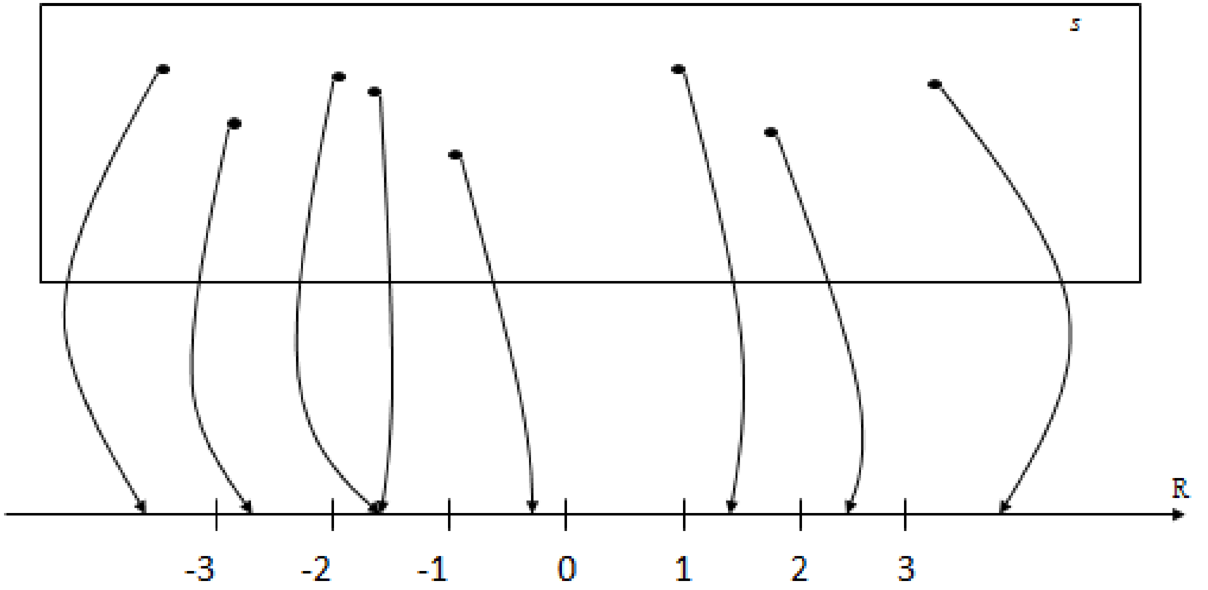
- Associated with each point s in the domain S the function X assigns one and only one value X(s) in the range R. (The set of possible values of X(s) is usually a proper subset of the real line; i.e., not all real numbers need occur. If S is a finite set with m elements, then X(s) can assume at most m different values as s varies in S.)

RV in graphical representation



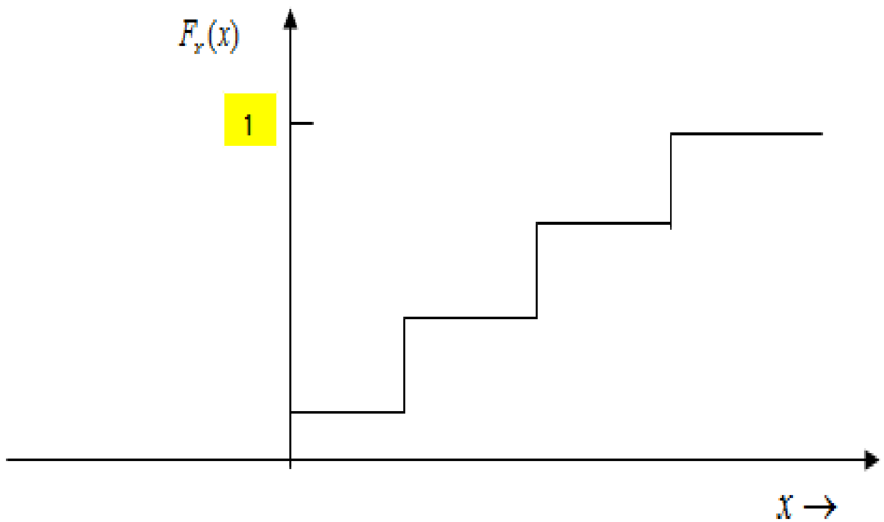
RV in graphical representation

- Random variable
  - A numerical value to each outcome of a particular experiment



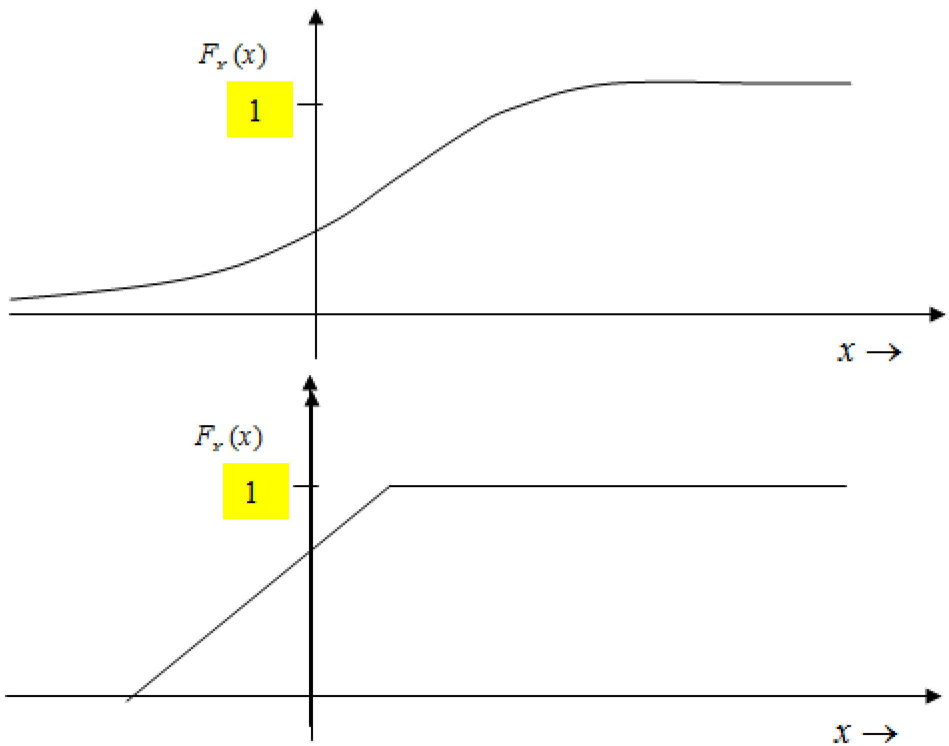
Discrete random variable

- A random variable is called a discrete random variable is piece-wise constant. Thus is flat except at the points of jump discontinuity. If the sample space is discrete the random variable defined on it is always discrete.



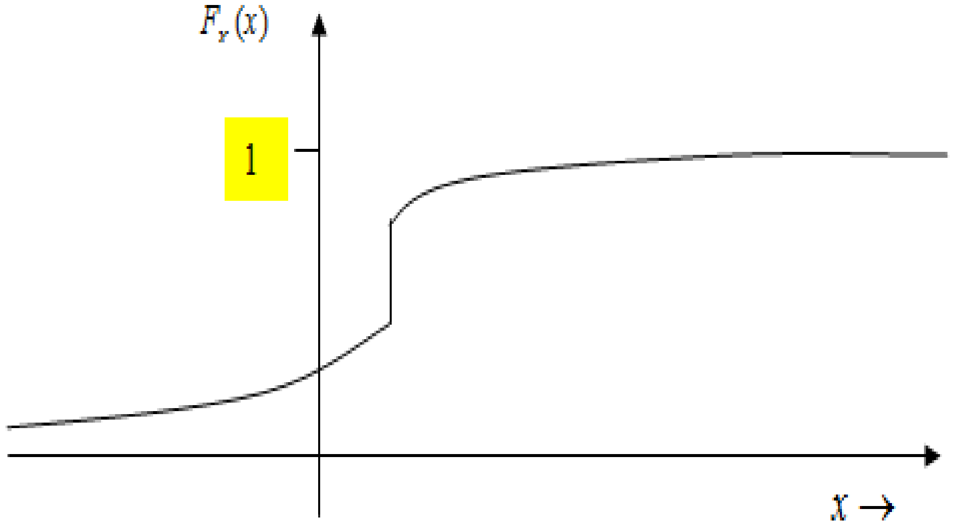
Continuous random variable

- X is called a continuous random variable if is an absolutely continuous function of x. Thus is continuous everywhere on and exists everywhere except at finite or countable infinite points.



Mixed random variable

- X is called a mixed random variable if has jump discontinuity at countable number of points and it increases continuously at least at one interval of values of x. For a such type RV X.



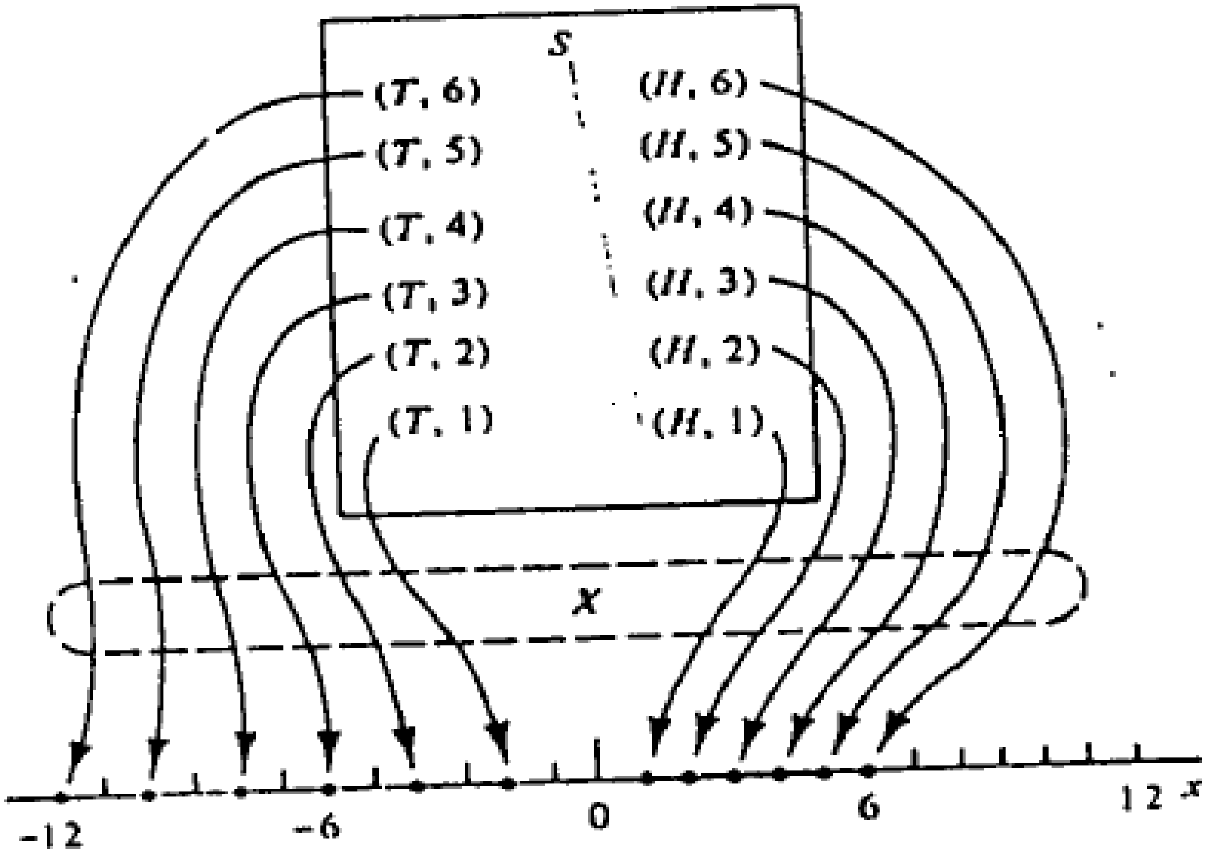
Random Variable

Review of the concepts

1. **Random Experiment**
2. **Random Event**
3. **Outcomes**
4. **Sample Space**
5. **Random Variable:**

   Mapping of sample space to a real line

Random Variable



Mapping of sample space to a real line

Distribution function

Probability Distribution Function

The probability $P(X \leq x)$ is the probability of the event $\{X \leq x\}$. i,e

$$F_x(x) = P\{X \leq x\}, \quad -\infty \leq x \leq \infty$$

Properties of CDF

The properties of a distribution function:

- $F_x(-\infty) = 0$
- $F_x(\infty) = 1$
- $0 \leq F_x(x) \leq 1$
- $F_x(x_1) \leq F_x(x_2)$, if $x_1 < x_2$ (Non-decreasing function)
- $P\{x_1 < X < x_2\} = F_x(x_2) - F_x(x_1)$
- $F_x(x^+) = F_x(x)$ (Continuous from the right)

Properties of CDF (contd..)

Proof for $F_x(x_2) - F_x(x_1)$

- The events $\{X \leq x_1\}$ and $\{x_1 < X < x_2\}$ are mutually exclusive, i.e. $\{X < x_2\} = \{X \leq x_1\} \cup \{x_1 < X < x_2\}$
- $P\{X < x_2\} = P\{X \leq x_1\} + P\{x_1 < X < x_2\}$
- $P\{x_1 < X < x_2\} = P\{X < x_2\} - P\{X \leq x_1\}$

$$= F_x(x_2) - F_x(x_1)$$

Properties of CDF (contd..)

If X is a discrete random variable taking values $x_i, \quad i = 1,2, \ldots, N$, then $F_x(x)$ must have a staircase function given by

$$F_x(x) = \sum_{i=1}^{N} P\{X = x_i\}\, u(x - x_i)$$

$$= \sum_{i=1}^{N} P(x_i)\, u(x - x_i)$$

where $u(.)$ is the unit-step function defined by:

$$u(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

If N is infinite, then

$$P(x_i) = P\{X = x_i\}$$

Probability density function

Probability Density Function

The probability density function of the random variable X is defined as the derivative of the distribution function:

$$f_x(x) = \frac{dF_x(x)}{dx}$$

Probability density function (contd..)

1. If the derivative of $F_x(x)$ exists then $f_x(x)$ exists
2. There may be places where $\frac{dF_x(x)}{dx}$ is not defined at points of abrupt change, then we shall assume that the number of points where $F_x(x)$ is not differentiable is countable.
3. For discrete random variables having a stair step form of distribution function.

$$f_x(x) = \sum_{i=1}^{N} P(x_1)\delta(x - x_1)$$

Properties of PDF

Properties of Density Functions.

- $0 \leq f_x(x)$ all x
- $\int_{-\infty}^{\infty} f_x(x)\,dx = 1$
- $F_x(x) = \int_{-\infty}^{x} f_x(x)\,dx = 1$
- $P\{x_1 < X < x_2\} = \int_{x_1}^{x_2} f_x(x)\,dx$

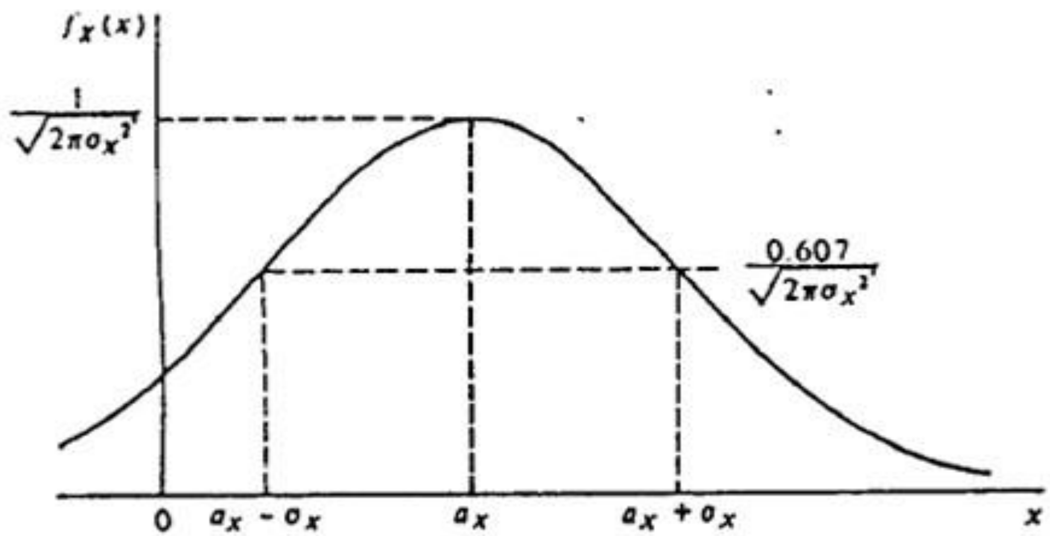
Gaussian Probability density function

Gaussian Density Function

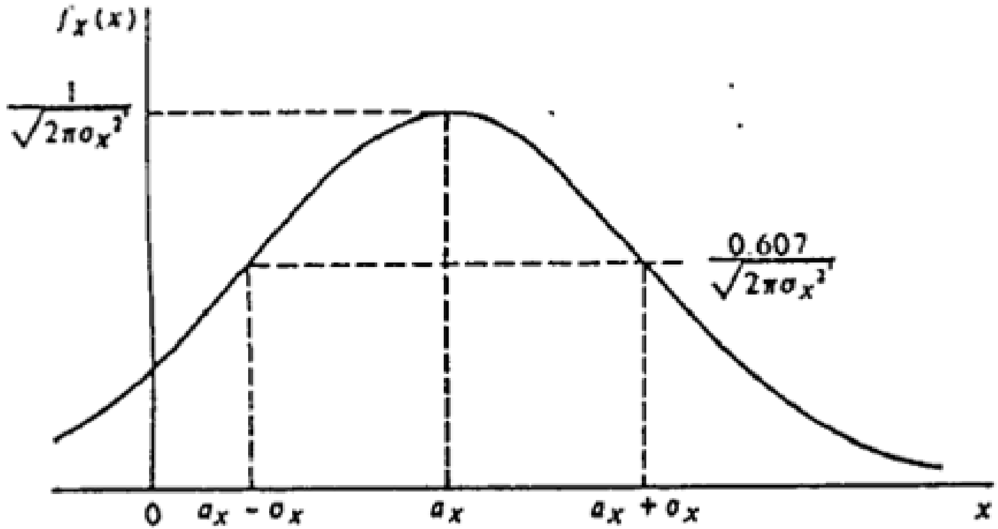
A random variable X is called Gaussian if its density function has the form

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-(x-a_x)^2/2\sigma x^2}$$

Where $\sigma_x > 0$ and $-\infty < a_x < \infty$ are real constants.



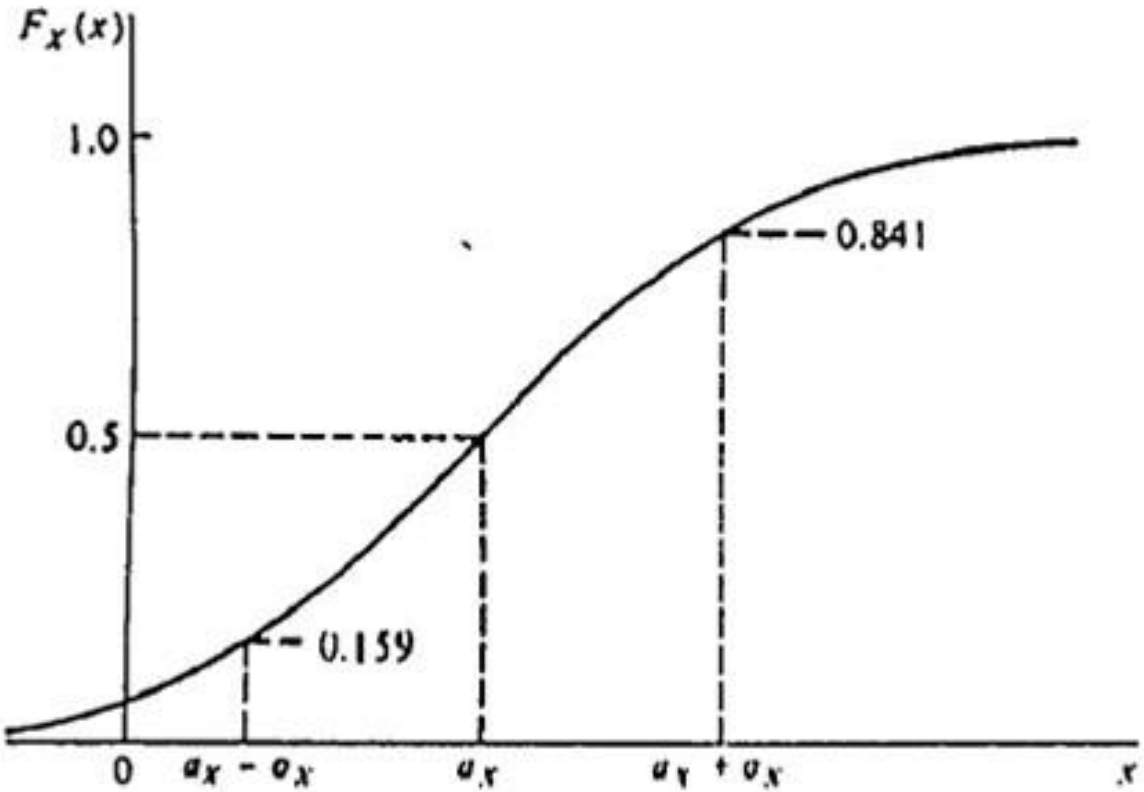
Gaussian Probability density function (contd..)



1. Its maximum value $(2\pi\sigma_2^x)^{-\frac{1}{2}}$ occurs at $x = a_x$.
2. Its “spread” about the point $x = a_x$ is related to σ_x.
3. The function decreases to 0.607 times its maximum at $x = a_x + \sigma_x$ and $x = a_x - \sigma_x$.
4. The Gaussian density is the most important of all densities. It enters into nearly all areas of engineering

Gaussian Probability density function (contd..)

$$F_x(x) = \frac{1}{\sqrt{2\pi\sigma_x^2}} \int_{-\infty}^{x} e^{-(\xi - ex)^2/2\sigma x^2} d\,\xi$$



Gaussian Probability density function (contd..)

- This integral has no known closed-form solution and must be evaluated by numerical methods.
- We could develop a set of tables of $F_x(x)$ for various x and α_x and σ_x as parameters (infinite number of tables).
- Only one table of $F_x(x)$ for the normalized (specific) values $\alpha_x = 0$ and $\sigma_{x=1}$ given by

$$F(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\xi^2/2} d\xi$$

which is a function of x only & tabulated for $x \geq 0$.

- For negative values of x we have

$$F(-x) = 1 - F(x)$$

Gaussian Probability density function (contd..)

- Making the variable change $u = \frac{\xi - a_x}{\sigma_x}$, we get

$$F_x(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{(x-a_x)/\sigma_x} e^{-ut/2}\, du = F\left(\frac{x - a_x}{\sigma_x}\right)$$

Binomial Probability density function

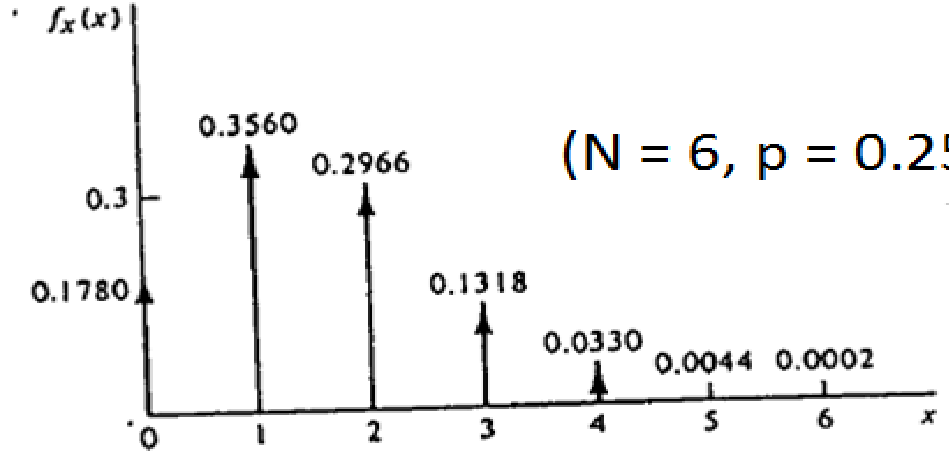
Binomial Density Function

$$f_x(x) = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta(x-k)$$

where $\binom{N}{k}$ is the binomial coefficient defined as

$$\binom{N}{k} = \frac{N!}{k!\,(N-k)!}$$

and $0 < p < 1,\ N = 1,2 \ldots \ldots \ldots .$



(N = 6, p = 0.25)

Binomial Probability density function (contd..)

1. The binomial density is applied to Bernoulli trail experiment, having only two possible outcomes on any given trial.
2. It applies to many games of chance, detection problems in radar and sonar, and many experiments

Binomial Probability density function (contd..)

By integration, the binomial distribution function is found:

$$F_X(x) = \sum_{k=0}^{N} \binom{N}{k} p^k (1-p)^{N-k} \delta(x-k)$$

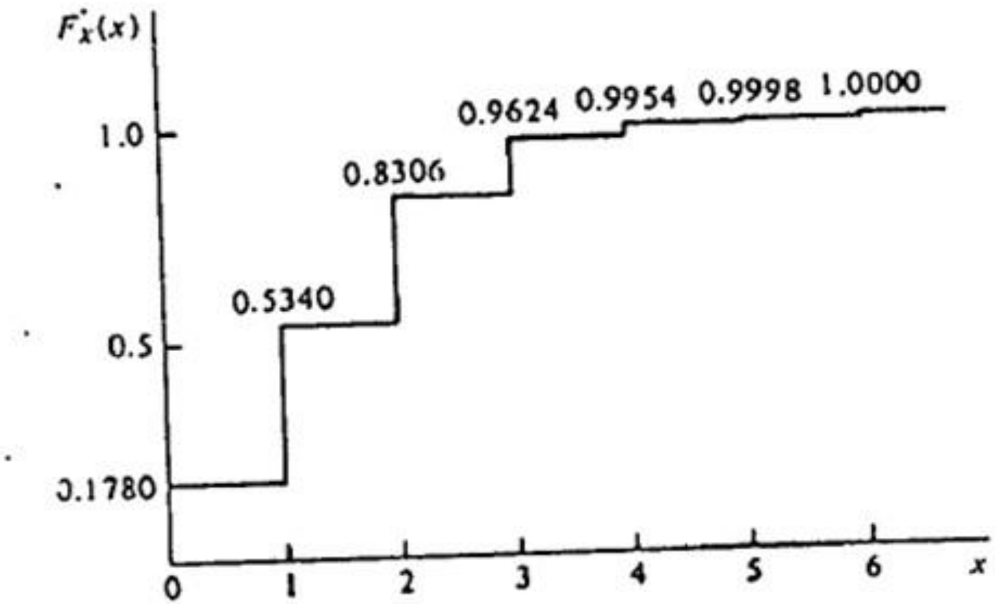


Figure: Binomial distribution function (N = 6, p = 0.25)

Poisson Probability density function

Poisson Density Function

$$f_X(x) = e^{-b} \sum_{k=0}^{\infty} \frac{b^k}{k!} \delta(x-k)$$

$$F_X(x) = e^{-b} \sum_{k=0}^{\infty} \frac{b^k}{k!} u(x-k)$$

where $b > 0$ is a real constant.

- These functions appear quite similar to binomial
- If $N \to \infty$ and $p \to 0$ for the binomial case in such a way that $N_p = b$, a constant, the Poisson case results.
- The Poisson random variable applies to a wide variety of counting-type applications.

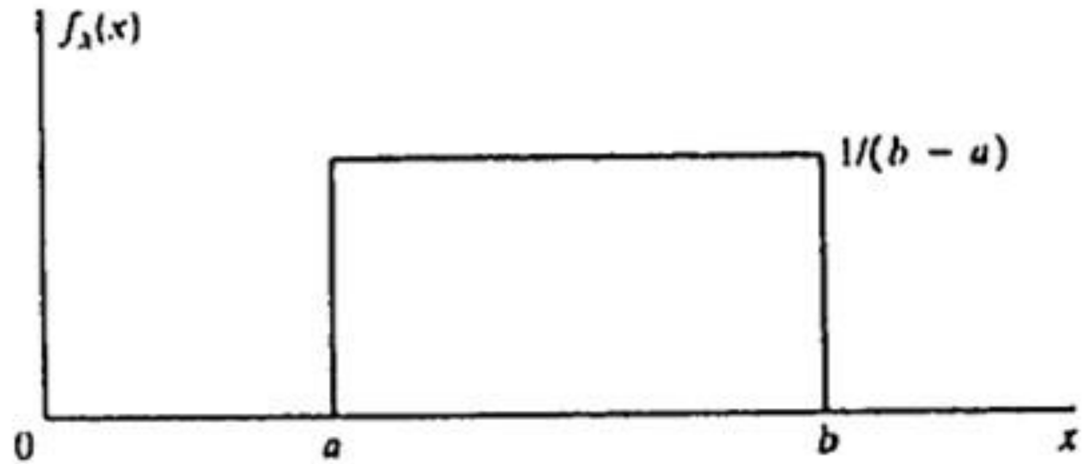
Poisson Probability density function (contd..)

- It describes
  - the number of defective units in a sample taken from a production line,
  - the number of telephone calls during a period of time,
  - the number of electrons emitted from a small section of a cathode in a given time interval, etc.
  - If the time interval of interest has duration T, and the events being counted are known to occur at an average rate λ and have a Poisson distribution, then $b = \lambda T$

Uniform Probability density function

Uniform Density Function

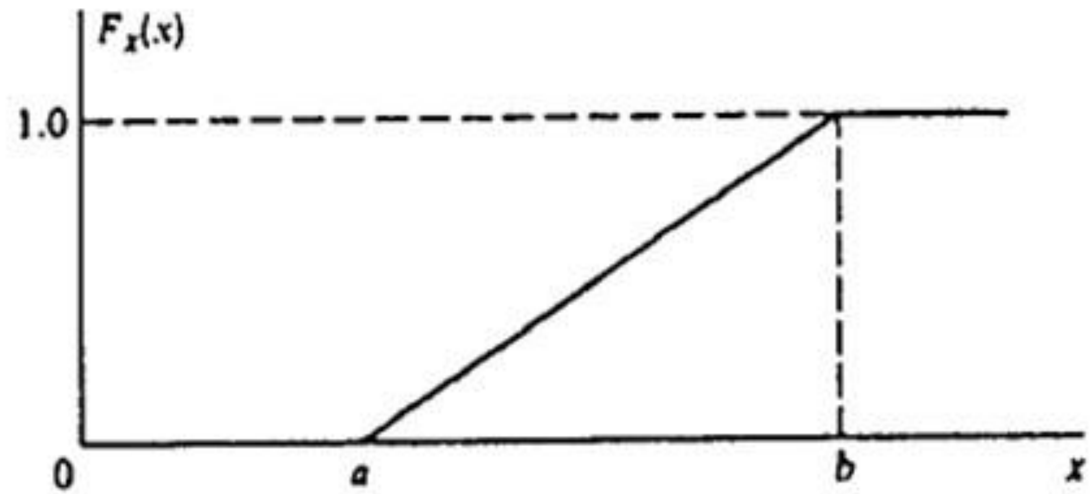
$$f_X(x) = \begin{cases} \dfrac{1}{b-a}, & a \le x \le b \\ 0, & elsewhere \end{cases}$$



for real constants $-\infty < a < \infty$ and $b > a$.

Uniform Probability density function (contd..)

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



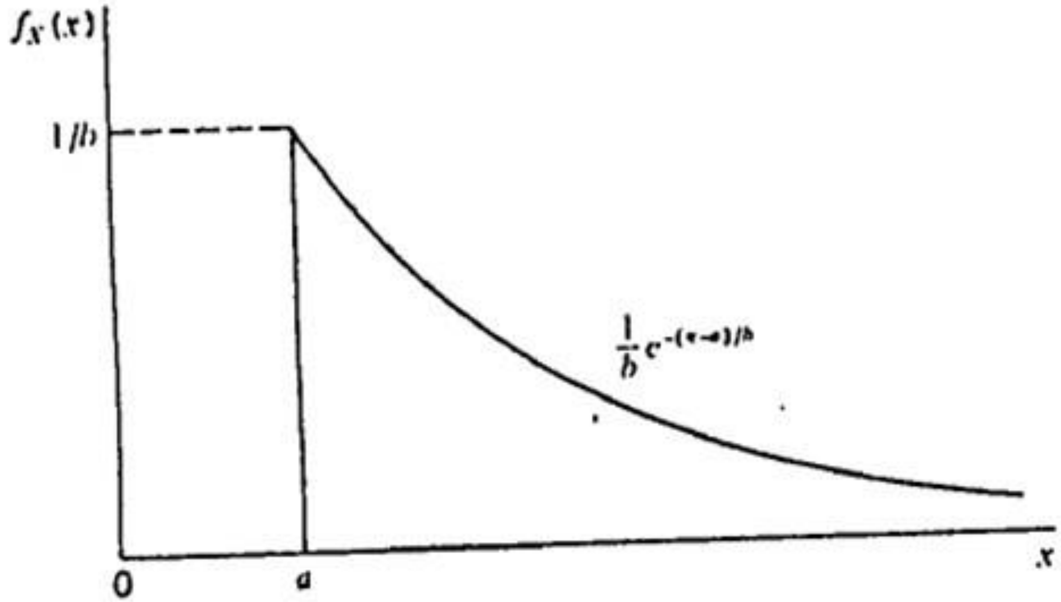
Uniform Probability density function (contd..)

- The error of quantization of signal samples prior to encoding in digital communication systems.
- Quantization amounts to “rounding off” the actual sample to the nearest of discrete quantum level.
- The quantization error introduced in the round-off process are uniformly distributed.

Exponential Probability density function

Exponential Density Function

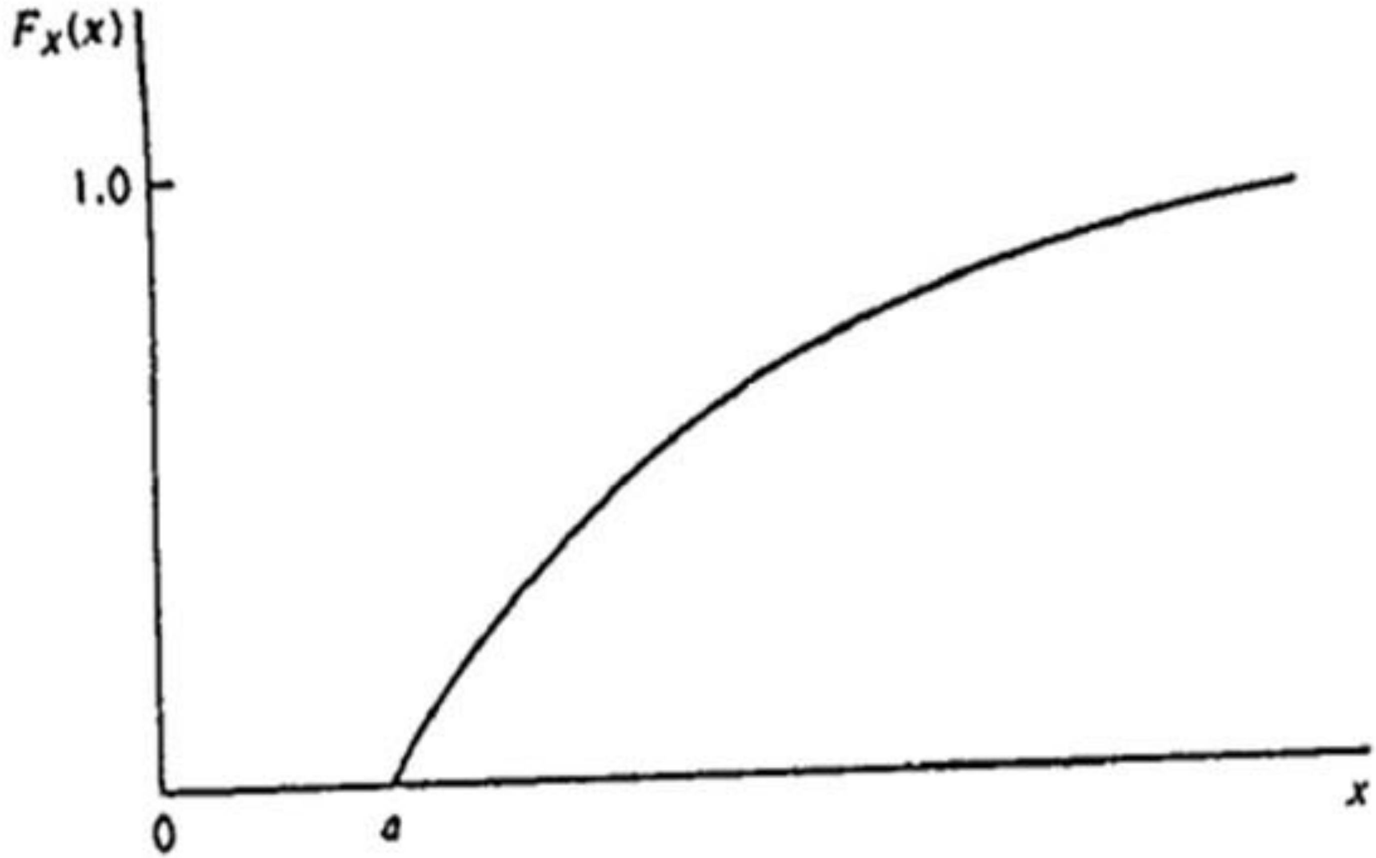
$$f_X(x) = \begin{cases} \frac{1}{b} e^{-(x-a)/b} & x > a \\ 0, & x < a \end{cases}$$



for real numbers $-\infty < a < \infty$ and $b > 0$

Exponential Pdf (contd..)

$$F_X(x) = \begin{cases} 1 - e^{-(x-a)/b} & x > a \\ 0, & x < a \end{cases}$$



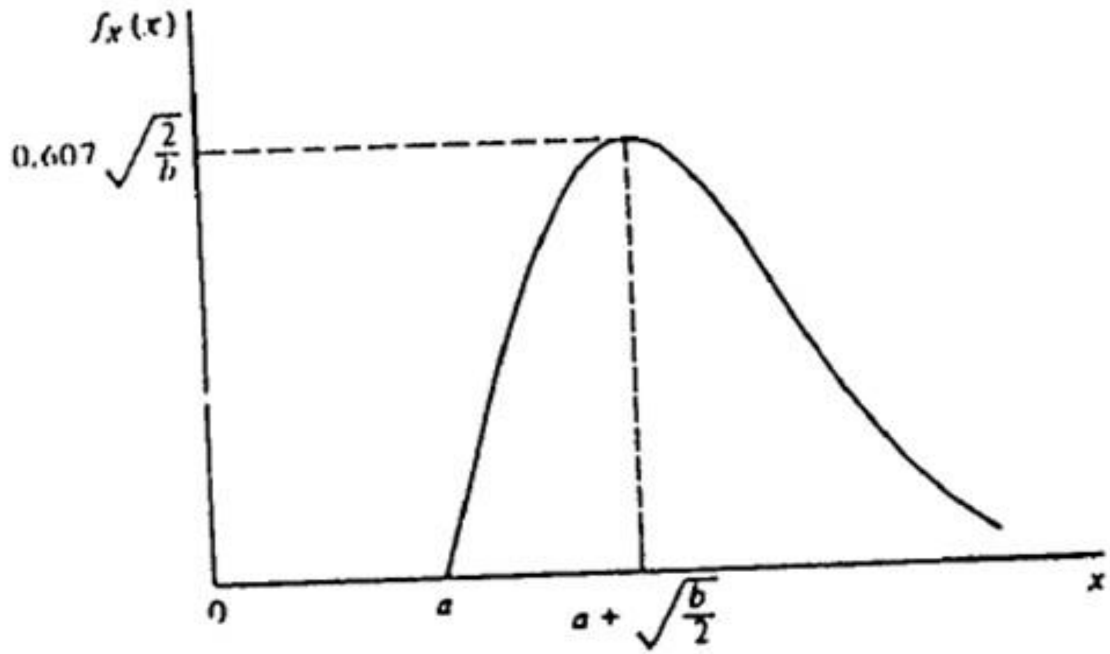
Exponential Pdf (contd..)

- The exponential density is useful in describing raindrop sizes when a large number of rainstorm measurements are made.
- It is also known to approximately describe the fluctuations in signal strength received by radar from certain types of aircraft.

Rayleigh Probability density function

Rayleigh Density Function

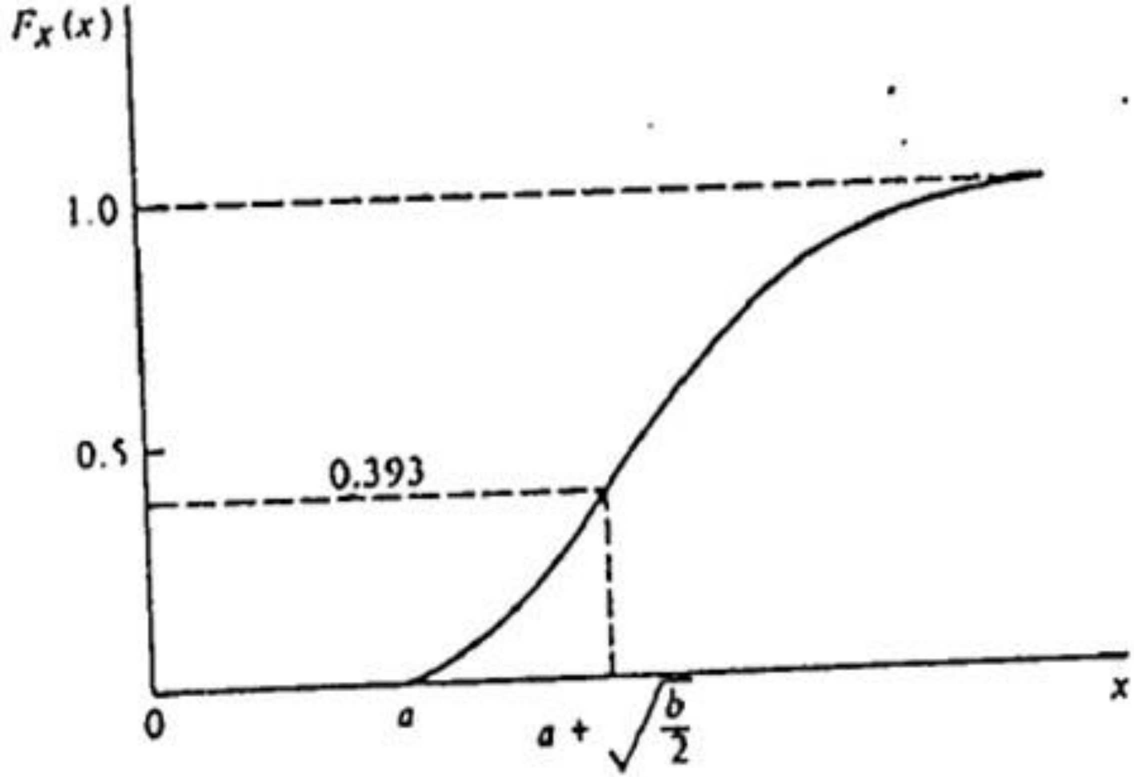
$$f_X(x) = \begin{cases} \frac{2}{b}(x-a)e^{-(x-a)2/b}, & x \geq a \\ 0, & x < a \end{cases}$$



for the real constants $-\infty < a < \infty$ and $b > 0$

Rayleigh Probability density function (contd..)

$$F_x(x) = \begin{cases} 1 - e^{-(x-a)2/b} & x \geq a \\ 0, & x < a \end{cases}$$



Rayleigh Probability density function (contd..)

- The Rayleigh density describes the envelope of white gaussian noise when passed through a band-pass filter.
- It is also is important in analysis of errors in various measurement systems.

Conditional distribution function

Conditional Distribution Function

- Let A and B be the two events & $P(B) \neq 0$, then

$$P(A|B) = \frac{P\{A \cap B\}}{P(B)}$$

- Let A be defined as the event $\{X \leq x\}$ for the random variable X.
- The resulting probability $P\{X \leq x|B\}$ is defined as the conditional distribution function of X , which is denote d by

$$F_X(x|B) = P\{X \leq x|B\} = \frac{P\{X \leq x \cap B\}}{P(B)}$$

where $\{X \leq x \cap B\}$ is the joint event $\{X \leq x\} \cap B$. This joint event consists of all outcomes ***s*** such that $X(s) \leq x$ and ***s*** ϵ B

Properties of Conditional distribution function

Properties of Conditional Distribution Function

- $F_x(-\infty|B) = 0$
- $F_x(\infty|B) = 1$
- $0 \leq F_x(\infty|B) \leq 1$
- $F_x(x_1|B) \leq F_x(x_2|B) \quad if\ x_1 < x_2$
- $P\{x_1 < X \leq x_2|B\} = F_x(x_2|B) - F_x(x_1|B)$
- $F_x(x^+|B) = F_x(x|B)$

Conditional density function

Conditional Density Function

The conditional density function of the random variable X is defined as the derivative of the conditional distribution function, and is given by

$$f_X(x|B) = \frac{dF_X(x|B)}{dx}$$

If $F_X(x|B)$ contains step discontinuities (when X is a discrete or mixed random variable), we assume that impulse functions are present in $f_X(x|B)$ to account for the derivatives at the discontinuities.

Properties of Conditional density function

Properties of Conditional Density Function

- $f_X(x|B) \geq 0$
- $\int_{-\infty}^{\infty} f_X(x|B)dx = 1$
- $F_X(x|B) = \int_{-\infty}^{x} f_X(\xi|B)\, d\xi$
- $P\{x_1 < X < x_1|B\} = \int_{x_1}^{x_2} f_X(x|B)dx$

Methods of conditioning event

Methods of Defining Conditioning Event

If event B is defined in terms of the random variable X as $B = \{X \leq b\}$, where b is some real number $-\infty < b < \infty$ & $P\{X \leq b\} \neq 0$, then we have

$$F_X(x|B) = P\{X \leq x|B\}$$

$$= P\{X \leq x|X \leq b\}$$

$$= \frac{P\{X \leq x \cap X \leq b\}}{P\{X \leq b\}}$$

Methods of conditioning event (contd..)

Case (i):

If $b \leq x$, then the event $\{X \leq b\}$ is an subset of the event $\{X \leq x\}$, so $\{X \leq x\} \cap \{X \leq b\} = \{X \leq b\}$. Then we have

$$F_X(x|X \leq b) = \frac{P\{X \leq x \cap X \leq b\}}{P\{X \leq b\}}$$

$$= \frac{P\{X \leq b\}}{P\{X \leq b\}} = 1 \qquad x \geq b$$

Methods of conditioning event (contd..)

Case (ii):

$$F_X(x|X \leq b) = \frac{P\{X \leq x \cap X \leq b\}}{P\{X \leq b\}}$$

$$= \frac{P\{X \leq x\}}{P\{X \leq b\}} = \frac{F_X(x)}{F_X(b)} \qquad x < b$$

Methods of conditioning event (contd..)

By combining the last two expressions, we have

$$F_X(x|X \leq b) = \begin{cases} \dfrac{F_X(x)}{F_X(b)} & x < b \\ 1, & x \geq b \end{cases}$$

From our assumption that the conditioning event has nonzero probability, $0 < F_X(b) \leq 1$, so the conditional distribution function is never smaller than the ordinary distribution function $F_X(x|X \leq b) \geq F_X(x)$

Methods of conditioning event (contd..)

Similarly the conditional density function is

$$f_X(x|X \le b) = \begin{cases} \dfrac{f_X(x)}{F_X(x)} = \dfrac{f_X(x)}{\int_{-\infty}^{b} f_X(x)dx} & x < b \\ 0, & x \ge b \end{cases}$$

From our assumption $0 < f_X(x) \le 1$, so the conditional density function is never smaller than the ordinary density function

$$f_X(x|X \le b) \ge f_X(x) \qquad x < b$$

The result can be extended to more general event

$B = \{a < X \le b\}$

UNIT-2
MOMENTS OF RANDOM VARIABLE