SYLLABUS COVERED

**Probability:** Axiomatic definition of probability and properties, conditional probability, multiplication rule. Theorem of total probability. Bayes’ theorem and independence of events.

**Random Variables:** Probability mass function, probability density function and cumulative distribution functions, distribution of a function of a random variable. Mathematical expectation, moments and moment generating function. Chebyshev's inequality.

**Standard Distributions:** Binomial, negative binomial, geometric, Poisson, hypergeometric, uniform, exponential, gamma, beta and normal distributions. Poisson and normal approximations of a binomial distribution.


**Limit Theorems:** Weak law of large numbers. Central limit theorem (i.i.d.with finite variance case only).
<table>
<thead>
<tr>
<th>Table of Contents</th>
<th>Page No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classical &amp; Axiomatic Probability</td>
<td>3 – 43</td>
</tr>
<tr>
<td>Geometric Probability</td>
<td>44 – 48</td>
</tr>
<tr>
<td>Random Variables</td>
<td>49 – 78</td>
</tr>
<tr>
<td>Mathematical Expectation</td>
<td>79 – 108</td>
</tr>
<tr>
<td>Moment Generating Function</td>
<td>109 – 118</td>
</tr>
<tr>
<td>Probability Inequalities</td>
<td>119 – 128</td>
</tr>
<tr>
<td>Standard Discrete Distributions</td>
<td>129 – 184</td>
</tr>
<tr>
<td>Standard Continuous Distributions</td>
<td>185 – 234</td>
</tr>
<tr>
<td>Joint Distributions of Two Random Variables</td>
<td>235 – 260</td>
</tr>
<tr>
<td>Bivariate Normal Distribution</td>
<td>261 – 278</td>
</tr>
<tr>
<td>Limit Theorems</td>
<td>279 – 300</td>
</tr>
</tbody>
</table>
Meanings of Probability:— It’s a measure of chance of occurrence of a phenomenon.

The word ‘Probability’ may be used to mean ‘the degree of belief’ of a person making a statement or proposition. It is used in the sense when we say that a certain football team will be the champion in a league or we say that the ‘Mahabharat’ is very probably the cooed of several Vachanas.

On the other hand, the word has a different meaning, when we use it in the context of an experiment that can be repeated any no. of times under identical conditions. By the probability of any outcome of the experiment we shall now mean the long run relative frequency of any particular outcome of the experiment. We use the probability in this sense when we say that the probability of getting a ‘head’ in tossing a coin is 3/4 or the probability that an article produced by a machine will defective is negligible. In statistics, we generally use the term in 2nd sense.

In probability and statistics, we concern ourselves to same special type of experiment.

Random Experiment:—
A random experiment or statistical experiment is an experiment in which—

(i) all possible outcomes of the experiment are known in advance, an outcome

(ii) any performance of the experiment results in a that is not known in advance.

(iii) The experiment can be repeated under identical or similar condition.

Ex: Consider an experiment of tossing a coin. If the coin does not stand on the side there are two possible outcomes: Head (H), Tail (T). On any performance of the experiment, one does not know what the result will be. Coin can be tossed as many times as desired under identical or similar condition. Hence, tossing of one is a random experiment.
(2) **Sample Space** — The collection of all possible outcomes of a random experiment is called the sample space of the random experiment. It is denoted by \( \Omega \) (or \( S \)). The elements of the sample space (\( \Omega \)) are called the 'sample points'.

**Ex: (1)** Consider a random experiment of 'tossing a coin' twice. Write down the sample space.

**Sol:** The sample space is — \( \Omega = \{HH, HT, TH, TT\} \)

The sample points are — HT, HT, TH, TT.

**Ex: (2)**

In each of the following experiment, what is the sample space?

i) a coin is tossed thrice.

ii) a die is rolled twice.

iii) a coin is tossed until a head appears.

**Sol:**

i) \( \Omega = \{HHH, HHT, HTT, HTH, TTH, THT, TTT\} \)

ii) \( \Omega = \{(i,j) \mid i,j = 1(1)6\} \) [arithmetic progression \(a(64)\)]

iii) \( \Omega = \{H, TH, TTH, TTT\} \)

**Ex: (3)** In each of the following experiments, what is the sample space?

i) In a survey of families with 3 children, the genders of the children are recorded in increasing order of their age.

**Sol:**

\( \Omega = \{BBB, BBG, BGB, GBB, GGB, GBB, GGB, GGG\} \)

ii) The experiment consists of selecting four items from a manufacturing output and observing whether or not each item is defective.

**Sol:**

\( \Omega = \{(a, b, c, d) \mid a, b, c, d \text{ is either defective or non-defective} \} \)

iii) Two cards are drawn from an ordinary deck of cards

(a) with replacement;

(b) without replacement

**Sol:**

(a) \( \Omega = \{ (x, y) \mid x, y = 1(1)52 \} \) [consisting 52 \(^2\) sample points]

(b) \( \Omega = \{ (x, y) \mid x, y = 1(1)52 \text{ but } x \neq y \} \) [consisting 52 \times 51 sample points]
Ex. (1) In each of the following experiments what is the sample space?

(i) Noting the lifetime of an electronic bulb.

(ii) A point is selected from a rod of unit length.

\[ \Omega = \{ \alpha \in \mathbb{R} : 0 < \alpha < \infty \} \quad \text{[Continuous sample space]} \]

\[ \Omega = \{ \alpha \in \mathbb{R} : 0 \leq \alpha \leq 1 \} \quad \text{[Here } \alpha \text{ is the distance of the selected point from the origin]} \]

(3) Trial: A trial refers to a special type of experiment in which there are two possible outcomes — 'success' and 'failure' with varying probability of success.

(4) Outcome: Result of an experiment.

(5) Sample: It is a part of the population and is supposed to represent the characteristic of the population.

(6) Event: An event is a subset of sample space.

(i) Elementary Event: If an event contains only one sample point, it's known as an elementary event.

(ii) Composite Event: If an event contains more than one sample point, it's known as a composite event.

Ex. (4) Consider the random experiment of tossing a fair coin twice. Identify elementary & composite events.

\[ \Omega = \{ HH, HT, TH, TT \} \]

The event (i) at least one head is \( A = \{ HH, HT, TH \} \), is called a composite event.

(ii) no head is \( B = \{ TT \} \), is called an elementary event.

Ex. (2) A club has 5 members A, B, C, D, E. It's required to select a chairman and a secretary. Assuming that 1 member can't occupy both positions. Write the sample space associated with this section. What's the event that member A is an officeholder.

\[ \Omega = \{ (x, y) : x, y \in \{ A, B, C, D, E \} \} \]

Here x stands for chairman and y stands for secretary.

Event is, \( P = \{ AB, BA, AC, CA, AD, DA, AE, EA \} \)

\[ \{ (x, y) : \text{If } x = A \text{ then } y = B, C, D, E. \text{ If } y = A \text{ then } x = B, C, D, E \} \]
Mutually Exclusive Events $\rightarrow$ Several events $A_1, A_2, \ldots$, An in relation to a random experiment, are said to be mutually exclusive (or disjoint) if any two of them cannot occur simultaneously; every time the experiment is performed is $A_i \cap A_j = \emptyset$, $\forall (i \neq j), 1 \leq i, j \leq n$.

Exhaustive Events $\rightarrow$ Several events $A_1, A_2, \ldots$, An in relation to a random experiment are said to be exhaustive events if any of them must necessarily occur, every time the experiment is performed that is $\bigcup_{i=1}^{n} A_i = \Omega$.

Equally Likely Cases (or events) $\rightarrow$ Several cases $A_1, A_2, A_3, \ldots$ are said to be equally likely if, after taking into consideration all relevant evidence, there is no reason to believe that one is more likely than the other.

Ex: $\rightarrow$ For a random experiment of ‘tossing a coin twice’, the sample space is $\Omega = \{HH, HT, TH, TT\}$.

Let $A$ be the event of getting at least one head and $B$ be the event of getting at most one head.

Then $A = \{HT, TH, HH\}$, $B = \{HT, TH, TT\}$.

$A \cup B = \Omega$ and $A \cap B = \emptyset$.

Hence, the event $A$ and $B$ are exhaustive but not mutually exclusive.

Let $C$ be the event of getting ‘no head’, then $C = \{TT\}$, $A \cup C = \Omega$, $A \cap C = \emptyset$.

Hence, the event $A$ and $C$ are exhaustive and mutually exclusive too.

The Classical Definition of Probability $\rightarrow$ If a random experiment can result in $N$ (finite) mutually exclusive, exhaustive and equally likely cases and $N(A)$ of them are favorable to the occurrence of the event $A$; then the probability of occurrence of $A$ is $P[A] = \frac{N(A)}{N}$.
Remarks: 1) Since $0 \leq N(A) \leq N$,
   
   $0 \leq P(A) \leq 1$.

2) By classical definition of probability of an event is a rational number between 0 and 1. But in general probability is a real no. between 0 and 1.

3) $P[A^c] = \frac{N-N(A)}{N} = 1 - \frac{N(A)}{N} = 1 - P(A)$.

Example:

1) A fair coin is tossed 3 times, what's the prob. of getting 'exactly 2 heads'.

2) What's the prob. of getting 'at least on tail'?

Solution:

1) $\Omega = \{HHH, HTH, THT, TTH, HTT, THH, TTT\}$

Since the coin is fair, $N=8$, elementary cases are equally likely. The events of getting two heads is $A = \{HHT, HTT\}$. Hence the no. of favorable cases $N(A) = 2$.

By classical definition $P[A] = \frac{N(A)}{N} = \frac{2}{8}$.

2) The event of getting 'at least one tail' is $N(\emptyset) = \Omega - \{HHH\} = 8 - 1 = 7$.

\[ \therefore \text{By classical definition, } P[\emptyset] = \frac{N(\emptyset)}{N} = \frac{7}{8}. \]

Limitation of Classical Definition:

1) It is assumed here that all the cases are equally likely. This def. of probability is found useful when applied to the outcomes of the games of chance. If the outcomes of a random experiment are not equally likely then this def. is not applicable.

2) This def. breaks down if the no. of all possible cases is infinite.

3) In real life it is not easy to identify the outcomes as equally likely.
Statistical or Empirical (Approach) Definition of Probability: Suppose A is an event of a random experiment. Suppose it is possible to repeat the experiment a large number of times under essentially similar condition.

Denote by n(A), the number of occurrence of A in ‘n’ repetitions. n(A) is called the frequency of A and \( \frac{n(A)}{n} \) is the relative frequency. A kind of regularity is observed when a large number of repetition is considered. It is an observed fact that the relative frequencies stabilize to a certain value as ‘n’ becomes large. This tendency seems to be inherent in the nature of a random experiment and stability of relative frequencies for large values of n constitutes the basis of statistical theory or statistical definition of probability. This kind of regularity in a random experiment is known as statistical regularity.

The limiting value of \( \frac{n(A)}{n} \) as \( n \rightarrow \infty \), is called the prob. of A, provided the limit exists.

Definition: If a random experiment is repeated under essentially similar conditions then the limiting value of the relative frequency of an event A, as the trials become in definitely large, is called the probability of event A, provided the limit exists.

Consider the Question: If a coin is tossed, what is the probability that it will turn up head.

Ans: Examine the result of tosses given below:

<table>
<thead>
<tr>
<th>No. of times the coin is tossed (n)</th>
<th>1</th>
<th>10</th>
<th>100</th>
<th>1000</th>
<th>2000</th>
<th>3000</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of times the head turns up [N(A)]</td>
<td>0</td>
<td>6</td>
<td>61</td>
<td>605</td>
<td>1907</td>
<td>1718</td>
</tr>
</tbody>
</table>
Thus, we get the relative frequencies as:

\[
\frac{\text{5}}{\text{10}}, \quad \frac{\text{61}}{\text{100}}, \quad \ldots, \quad \frac{\text{605}}{\text{1000}}, \quad \ldots, \quad \frac{\text{1801}}{\text{2000}}, \quad \ldots, \quad \frac{\text{1718}}{\text{3000}}, \quad \ldots
\]

As the no. of tossing increases, the relative frequency tends to stabilize at 0.6. Therefore, the probability of getting a head in a tossing of a coin is 0.6.

**Remarks:** If in a random experiment all possible cases are not equally likely, then we can’t apply classical definition in this case. If the experiment can be repeated a large no. of times, then probability of an event A can be obtained by statistical definition, this is an improvement over the statistical defn.

**Limitations:**

1. If an experiment is repeated a number of times, the experimental conditions may not remain identical or homogeneous.
2. \[ \lim_{n \to \infty} \frac{n(A)}{n} \] may not be unique.

**Subjective Probability:** In everyday’s life, we hear or make statements such as “probably I shall miss the train”, “probably Mr.Raj will be at home now.” Such statements can be made more precise by “the chance of missing the train is 60%,” “the chance that Mr.Raj will be at home now is 75%” etc. Here 60%, 75%, etc. measures one’s belief in the occurrence of the event. This subjective method is another method of assigning probabilities of various events based on the personal beliefs.

When the experiment is not repeatable, this method may be adopted for assigning probabilities to events. Since, different persons may assign different probabilities, one can’t arrive at an objective conclusion using probabilities assigned by subjective methods.
Probability & Statistics: The problem in Probability is —
"Given a stochastic model what we can say about the outcome."
The problem in Statistics is —
"Given a sample what we can say about the population."

Set Theory
1. Point / Element
2. Set
3. Universal Set
4. Null set
5. A is a subset of B
6. A is a superset of B

Probability Theory
Elementary Event
Event
Sample Space
Impossible event
A implies B
A is implied by B

Ex. 1. Let A, B, C are 3 events. Then the expression of following events in set notations:
(i) Only A occurs: \( A \cap B^c \cap C^c \)
(ii) A occurs: \( A \)
(iii) Both A and B, but not C occur: \( A \cap B \cap C^c \)
(iv) All 3 events occur: \( A \cap B \cap C \)
(v) At least one occur: \( A \cup B \cup C \)
(vi) At least two occur: \( (A \cap B) \cup (B \cap C) \cup (C \cap A) \)
(vii) One and no more occur: \( (A \cap B^c \cap C^c) \cup (B \cap C \cap A^c) \cup (C \cap A \cap B^c) \)
(viii) Two and no more occur: \( (A \cap B \cap C^c) \cup (B \cap C \cap A^c) \cup (C \cap A \cap B^c) \)

Ex. 2. Eight students are arranged at random
(a) in a row and (b) in a column.
Find the probability that two given students will be next to each other.
Sol. (a) Req. prob. = \( \frac{7! \cdot 2!}{8!} \)
(b) Rev. prob. = \( \frac{6! \cdot 2!}{7!} \)

Ex. 3. The nine digits 1, 2, 3, ..., 9 are arranged in random order to form a nine-digit number. Find the prob. that 1, 2 and 3 appear as neighbours in the order mentioned.
Sol. Req. prob. = \( \frac{7!}{9!} = \frac{1}{72} \)

Ex. 4. Find the prob. that seven people have birthdays on 7 different days of the week, assuming equal prob. for the seven days.
Sol. Rev. prob. = \( \frac{7!}{7^7} \)
No. of distinguishable or distinct arrangement of \( n \) balls (objects) into \( n \) cells when

(I) balls are distinguishable and exclusion principle followed.

(II) " " " " but " " NOT " "

(III) " " indistinguishable and " " NOT followed.

(IV) " " " " but " " NOT " "

Exclusion Principle:— The principle of excluding a cell from taking more than one ball (object) while distributing \( n \) balls (objects) into \( n \) cells, i.e., to exclude one from an occupied cell which is occupied.

**CASE-I:** Let \( u(n,n) \) denotes the no. of distinguishable distributions of \( n \) balls into \( n \) cells.

Hence, \( u(n,n) = 0 \) if \( n > n \).

For \( n \leq n \), we have \( u(n,n) \)

\[
= \left( \text{no. of ways in which } \frac{1}{1} \text{ ball can be placed} \right) \times \left( \text{no. of ways in which } \frac{2}{2} \text{ ball can be placed} \right) \times \ldots \times \left( \text{no. of ways in which } \frac{n}{n} \text{ ball can be placed} \right)
\]

\[
= n(n-1) \ldots \ldots (n-n+1) = \binom{n}{n} n^n = n^n.
\]

**CASE-II:** Maxwell-Boltzmann Statistics

Here \( u(n,n) = n \cdot n \times \ldots \times n \times n = n^n. \)

**CASE-III:** Fermi-Dirac Statistics

Here \( u(n,n) = 0 \) for \( n > n \). For \( n \leq n \), \( u(n,n) = \frac{\binom{n}{n} n^n}{n!} = \binom{n}{n}. \)

**CASE-IV:** Bose-Einstein Statistics

\( u(n,n) = \) no. of distinguishable arrangements of \( n \) dots and \( (n-1) \) bars

\[
= \frac{(n+n-1)!}{n! (n-1)!} = \binom{n+n-1}{n}.
\]
Ex. 1. 2 cards are drawn from a well-shuffled cards. What's the probability that both extracted cards are aces.

Sol. Here, total no. of cases = no. of ways in which 2 cards can be drawn from 52 cards

No. of favourable cases = No. of ways of getting two ace from 4 aces

So, Required probability = \( \frac{\text{No. of favourable cases}}{\text{Total no. of cases}} = \frac{4 \times 3}{52 \times 51} = \frac{1}{221} \)

Ex. 2. Two dice are thrown n times in succession. What is the prob. of obtaining double 6 at least once. Also determine the minimum no. of throws required to accomplish the objective with a probability \( > \frac{1}{2} \).

Sol. (i) No. of throws resulted in with required probability a double six at least once

\[ \frac{\text{total no. of all possible cases}}{36^n - 35^n} = 1 - \left( \frac{35}{36} \right)^n = \phi_n, \text{ say} \]

\[ \phi_n > \frac{1}{2} \Rightarrow \left( \frac{35}{36} \right)^n < \frac{1}{2} \]

\[ \Rightarrow n \left( \log_{36} 35 - \log_{36} 36 \right) = -\log_2 \]

\[ \Rightarrow n < \frac{\log_2}{\log_{36} 35 - \log_{36} 35} \]

\[ \Rightarrow n \min = \frac{\log_2}{\log_{36} 35 - \log_{36} 35} \approx 24 \]

Ex. 3. A certain number n of distinguishable balls is distributed among N compartments. What is the prob. that a certain specified compartment will contain k balls?

Sol. Total no. of cases = No. of ways in which n distinguishable balls can be distributed among N compartments without following exclusion principle.

No. of favourable cases = \( \binom{n}{k} \times \binom{n-k}{N-1} \times \binom{N-1}{k-1} \) \( \times \binom{N-k-1}{N-1} \times \binom{N-k-1}{k-n} \)

\[ \Rightarrow \text{Reqd. prob.} = \frac{\binom{N}{k} \cdot (N-1)^{n-k}}{N^n} \]
Ex. 4. In an urn there are \( n \) groups of \( p \) objects in each. Objects in different groups are distinguished by some characteristic property. What is the prob. that among \((\alpha_1+\ldots+\alpha_n)\) objects taken, \([0 \leq \alpha_i \leq p \; i=1(1)n]\), there are \( \alpha_1 \) of one group, \( \alpha_2 \) from another group, \ldots and so on.

Sol. The total no. of cases = \[\binom{n p}{\alpha_1+\alpha_2+\ldots+\alpha_n}\]

Favourable cases = \[
\frac{\text{(no. of distinguishable arrangements of } \alpha_1, \alpha_2, \ldots, \alpha_n)}{\text{(no. of ways in which } \alpha_1 \text{ comes from one group, } \alpha_2 \text{ from 2nd and so on)}}
\]

= \[
\frac{n!}{1! 2! \ldots n!} \times \binom{p}{\alpha_1} \binom{p}{\alpha_2} \ldots \binom{p}{\alpha_n}
\]

Ex. 5. There are \( N \) tickets numbered \( 1, 2, \ldots, N \) of which \( n \) are taken at random in an increasing order of their numbers \( \alpha_1 < \alpha_2 < \ldots < \alpha_n \). What is the prob. that \( \alpha_m = M \).

Sol. The \( n \) tickets can be taken in \( \binom{N}{n} \) ways. We assume that these are equally likely.

In order that \( \alpha_m = M \), it is necessary & sufficient that \( \binom{n-1}{n-m} \) tickets have numbered from \( 1 \) to \( M-1 \). How many tickets have numbers from \( N-m \) to \( N \) and one ticket has the number \( M \). Hence, the no. of favourable cases are \( \binom{M-1}{n-m} \binom{N-M}{n-m} \).

So, Rev. prob. is = \[
\binom{N}{n} \frac{\binom{M-1}{n-m} \binom{N-M}{n-m}}{\binom{N}{n}}
\]

Ex. 6. An urn contains \( a \) 'white' and \( b \) 'black' balls. Balls are drawn one by one until only those of the same colour are left. What is the prob. that they \( a \) are white.

Sol. Let \( E \) be the given experiment and \( A \) be the desired event.
Let \( E' \) be the desired experiment of drawing all the balls one by one and \( A' \) the event that the last ball drawn is white. Then \( A \) happens in \( E \iff A' \) happens in \( E' \). Hence, \( P(A) = P(A') \). Since the ball are drawn at random in \( E' \), \( P(A') \) is also the prob. that the first ball drawn is white and hence is \( \frac{a}{a+b} \).
Ex. 7. Three numbers are chosen from the first 30 natural numbers. What's the prob. that the chosen number will be in (a) A.P. (b) G.P.

Solution:
(a) \( N = \{1, 2, \ldots, 30\} \)

Three numbers can be chosen from 30 natural numbers in \( \binom{30}{3} \) ways which are assumed to be equally likely.

In order that the three numbers will be of the form \( a, a+k, a+2k \quad (k \geq 1) \), \( a \) must satisfy \( 1 \leq a \leq 19 \) and \( m, m+k, m+2k \quad (k \geq 1) \), \( m \) must satisfy \( 1 \leq m \leq 30 - 2k \).

Hence, the number of such A.P.s is \( \sum_{k=1}^{19} \binom{30-2k}{3} \)

So, the required probability is \( \frac{\sum_{k=1}^{19} \binom{30-2k}{3}}{\binom{30}{3}} = 0.0517 \).

(b) We count the triplets (arranged in increasing order) whose terms form a G.P. by listing them as follows:

<table>
<thead>
<tr>
<th>Common ratio</th>
<th>Triplet</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>( {i, 2i, 4i}, 1 \leq i \leq 7 )</td>
</tr>
<tr>
<td>3</td>
<td>( {i, 3i, 9i}, 1 \leq i \leq 3 )</td>
</tr>
<tr>
<td>4</td>
<td>( {1, 4, 16} )</td>
</tr>
<tr>
<td>5</td>
<td>( {1, 5, 25} )</td>
</tr>
<tr>
<td>3/2</td>
<td>( {4, 8, 9}, {8, 12, 18}, {12, 18, 27} )</td>
</tr>
<tr>
<td>5/2</td>
<td>( {4, 10, 15} )</td>
</tr>
<tr>
<td>4/3</td>
<td>( {9, 12, 16} )</td>
</tr>
<tr>
<td>5/3</td>
<td>( {9, 15, 25} )</td>
</tr>
<tr>
<td>5/4</td>
<td>( {16, 20, 25} )</td>
</tr>
</tbody>
</table>

\( \therefore \) Hence the required prob. is \( \frac{19}{\binom{30}{3}} = 0.0517 \).
Axiomatic Approach

1) Explain the concept of Kolmogorov's Axiomatic definition of probability. Using this show that —
   i) \( P(\emptyset) = 0 \), when \( \emptyset \) is null set.
   ii) \( P(A) \leq 1 \), for any event \( A \).

Solu:—

Axiomatic Definition:— Let \( \Omega \) be the sample space of a random experiment and \( \mathcal{E} \) be a \( \mathfrak{F} \)-field of events of \( \Omega \). A set function \( P(\cdot) \) defined on \( \mathcal{E} \) is called a probability measure if it satisfies the following conditions:

Axiom I (Axiom of non-negativity): \( P(A) \geq 0 \ \forall A \in \mathcal{E} \).

Axiom II (Axiom of unit norm): \( P(\Omega) = 1 \).

Axiom III (Axiom of countable additivity): If \( A_i, i=1(1)n \) be a disjoint sequence of events in \( \mathcal{E} \), then
\[
P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i)
\]

In axiomatic approach probability is regarded as a set function.

i) Let \( A_1, A_2, \ldots \) be events in \( \mathcal{E} \). \( \exists A_i = \emptyset, \forall i \). Then
\[
\bigcup_{i=1}^{n} A_i = \emptyset \quad \text{and since} \quad A_i \cap A_j = \emptyset \cap \emptyset = \emptyset, \forall i \neq j.
\]

Then \( A_i's \) are also mutually exclusive (i.e. disjoint).

i) By the axiom of countable additivity, we have
\[
P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i)
\]

or, \( P(\emptyset) = P(\emptyset) + P(\emptyset) + P(\emptyset) + \ldots \ldots \)

But this can happen if either \( P(\emptyset) = 0 \) or \( P(\emptyset) = \infty \) or \( P(\emptyset) \) is not possible.

So, \( P(\emptyset) = 0 \). (Proved)

ii) As \( A \subset \Omega \) for each \( A \in \mathcal{E} \),
\[
\Rightarrow P(A) \leq P(\Omega).
\]
Now, from the axiom of unit norm, we know \( P(\Omega) = 1 \). So,
we get \( P(A) \leq 1 \) for any event \( A \).
\[ A \cup A^c = \Omega, \quad A \cap A^c = \emptyset, \quad A = A^c \]

by finite additivity of \( P(\cdot) \),

\[ P[A \cup A^c] = P[A] + P[A^c] - P[A \cap A^c] \]

\[ \therefore P(A) = P(A) + P(A^c) \]

\[ 0 \leq P(A^c) = 1 - P(A), \text{ by Axiom I} \]
\[ \therefore P[A] \leq 1. \]

2) (a) Let \( A_1, \ldots, A_n \) be \( n \) events \( \exists P(A_i) = 1 \).

\( \forall i = 1(1)n \).

(b) Let \( A_1, A_2, \ldots \) be the events \( \exists P(A_i) = 0, \forall i = 1, 2, \ldots \)

then show that \( P(\bigcup_{i=1}^{\infty} A_i) = 0 \).

(c) If the events \( A_i \)'s are mutually exclusive and exhaustive events of \( \Omega, i = 1, 2, \ldots \)

S.T. \( \sum_{i} P(A_i) = 1. \)

So (a) if \( A_i, i = 1(1)n \) be events in \( \mathcal{A} \), then Borel-Cantelli inequality gives

\[ P(\bigcap_{i=1}^{n} A_i) \geq \sum_{i=1}^{n} P(A_i) - n + 1 \quad (i) \]

From the axiom of unit norm, \( P(\Omega) = 1 \).

As \( A \subset \Omega, \forall A \in \mathcal{A} \).

\[ \therefore P(A) \leq P(\Omega) = 1 \quad \therefore P(A) \leq 1 \quad (ii) \]

Here, \( P(A_i) = 1, \forall i = 1(1)n \) \quad (iii)

So, from (i), (ii), (iii) we get \( P(\bigcap_{i=1}^{n} A_i) = 1. \)

(b) If \( P(A_i) = 0 \), we know from Boole's inequality

\[ P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i) \text{ and } P(A) > 0. \]

So, if \( P(A_i) = 0, \forall i \geq 1 \), we get \( P(\bigcup_{i=1}^{n} A_i) = 0. \)

Hence the result is proved.

(c) Since, \( A_i \)'s are exhaustive events, then \( \bigcup_{i} A_i = \Omega \)

\[ \therefore P(\bigcup_{i} A_i) = P(\Omega) = 1 \]

Again \( A_i \)'s are mutually exclusive, \( P(\bigcup_{i=1}^{n} A_i) = P(\bigcap_{i=1}^{n} A_i) \)

\[ \therefore P(\bigcap_{i} A_i) = 1 \quad \therefore \sum_{i} P(A_i) = 1. \]
Independence Of Events

i) Define mutually exclusive, exhaustive and mutually independent events. Let the two events be mutually exclusive are they mutually independent.

ii) Show by an example that pair wise independence does not necessarily imply mutual independence.

iii) Distinguish between pair wise and mutual independence of a finite set of events.

iv) Show that if \( A_1, A_2, A_3 \) are mutually independent then \( A_1^c, A_2^c, A_3^c \) are also mutually where \( A^c \) is the complement of \( A \).

Mutually exclusive events: Several events \( A_1, A_2, \ldots \)
--- An in relation to a random experiment are said to be mutually exclusive (or, disjoint) if any two of them can't occur simultaneously. Every time the experiment is performed is \( A_i \cap A_j = \emptyset \), \( i \neq j \), \( i, j = 1(\Omega) \).

Exhaustive Events: Several events in relation to a random experiment are said to be exhaustive events if at least one of them necessarily occurs. Thus the events \( A_1, A_2, \ldots, A_n \) are exhaustive if

\[
\bigcup_{i=1}^{n} A_i = \Omega.
\]

Pairwise Independence of a set of events:
A set of events \( \{A_1, A_2, \ldots, A_n\} \) is said to be pairwise independent if

\[
P(A_i \cap A_j) = P(A_i) P(A_j), \quad i \neq j, \quad i, j.
\]

Here we have \((\Omega)\) restrictions.

Mutually independence of a set of events:
A set of events \( \{A_1, A_2, \ldots, A_n\} \) is said to be mutually independent if

\[
P(A_i \cap A_j \cap \cdots \cap A_k) = P(A_i) P(A_j) \cdots P(A_k), \quad i \leq j \leq k.
\]

\[
P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P[A_i].
\]

i.e., \( P[\bigcap_{i=1}^{n} A_i] = \prod_{i=1}^{n} P[A_i] \).
The idea of mutual independent emerges from the following fact:
\[ P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) P(A_2) \ldots P(A_n) \]

Under statistical independence if all the conditional probabilities become equal to the respective unconditional probabilities, then we get:
\[ P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) P(A_2) \ldots P(A_n) \]
Here we have \((2^n - 1 - n)\) restrictions.

If two events are mutually exclusive then they will not be mutually independent.

A fair coin is tossed twice, \(S = \{HH, HT, TH, TT\}\)

- \(A\): Two heads appear, \(\{HH\}\).
- \(B\): One head & one tail appear, \(\{HT, TH\}\) ("exactly one head appears")

This two events are mutually exclusive.

\[ A \cap B = \emptyset \]

\[ \therefore P(A \cap B) = 0 \]

\[ P(A) = \frac{1}{4}; \quad P(B) = \frac{2}{4} = \frac{1}{2} \]

\[ \therefore P(A \cap B) \neq P(A) P(B) = \frac{1}{8} \]

They are not mutually independent.

**Note:** Mutually exclusive events in general are not independent and also, independent events are not in general mutually exclusive.

**ii & iii:**

**Distinction between Pairwise Independence and Mutually Independence:**

If \(A_1, A_2, \ldots, A_n\) are pairwise independent, then:
\[ P(A_i \cap A_j) = P(A_i) P(A_j) \quad \text{for } i, j \text{ (} i \neq j \text{)}, \]

but for mutually independence it is necessary that all of the \((2^n - p - 1)\) equations hold as mentioned earlier. It is evident that mutually independence implies pairwise independence but the converse may not be true. An example to show that pairwise independence does not imply mutually independence.
Suppose a fair coin is tossed twice.

Let $A$: the first toss gives a head.
$B$: the second toss gives a head.
$C$: both give the same outcome.

$S_2 = \{HH, HT, TH, TT\}$

$A = \{HH, HT\}$; $B \cap C = \{HH\}$

$B = \{HH, TH\}$; $A \cap C = \{HH\}$

$C = \{HT, TT\}$; $A \cap B \cap C = \{HH\}$

$A \cap B = \{HH\}$

$\therefore P(A \cap B) = \frac{1}{4} = P(A) P(B)$;

$P(B \cap C) = \frac{1}{4} = P(B) P(C)$; $P(A \cap C) = \frac{1}{4} = P(A) P(C)$.

$\therefore A, B, C$ are pairwise independent.

$\therefore P(A \cap B \cap C) = \frac{1}{4} \neq P(A) P(B) P(C)$.

$\therefore A, B, C$ are not mutually independent.

We know that mutually independence does necessarily imply pairwise independence.

So, $A_1, A_2, A_3$ are both mutually and pairwise independent.

$P(A_1 \cap A_2 \cap A_3) = P(A_1) P(A_2) P(A_3)$

$P(A_1 \cap A_2) = P(A_1) P(A_2)$.

$P(A_2 \cap A_3) = P(A_2) P(A_3)$;

$P(A_1 \cap A_3) = P(A_1) P(A_3)$.

$P(A^c_1 \cap A_2 \cap A_3) = 1 - P(A_1 \cup A_2 \cup A_3)$

$= 1 - [P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)]$

$= \{1 - P(A)\} \{1 - P(B)\} \{1 - P(C)\} \\
+ P(B) P(C) \{1 - P(A)\}$

$= \{1 - P(A)\} \{1 - P(B)\} \{1 - P(C)\} \\
+ \{1 - P(A)\} \{1 - P(A)\} \{1 - P(B)\}$

$= \{1 - P(A)\} \{1 - P(A)\} \{1 - P(A)\}$

$P(A_1^c) P(A_2^c) P(A_3^c)$

$\therefore A_1^c, A_2^c, A_3^c$ are mutually independent if $A_1, A_2, A_3$ are mutually independent.
In a sample space of 8 equally likely points, find the following:

i) Three events that are pairwise independent but not mutually independent.

so \( \text{consider a random experiment of a certain coin is thrown thrice.} \)

Sample space is, \( S = \{\text{HHH, HHT, HTH, TTH, HTT, THT, TTH, TTT}\} \)

and \( P(A) = \frac{1}{8} \) \( \forall A \in S \).

Define, \( A_1: \text{At least two heads} = \{\text{HHH, HHT, HTH, THT}\} \)

\( A_2: \{\text{HHH, HHT, HTH, THT}\} \)

\( A_3: \{\text{HHH, HHT, TTH, TTT}\} \)

\( P(A_1) = \frac{1}{2} \) \( \forall i = 1, 2, 3 \).

Now, \( P(A_1 \cap A_2) = P(\{\text{HHH, HHT, THT}\}) = \frac{3}{8} = \frac{1}{4} \)

\( P(A_2 \cap A_3) = P(\{\text{HHH, HHT, THT}\}) = \frac{1}{4} = P(A_1)P(A_2) \)

\( P(A_1 \cap A_2 \cap A_3) = P(\{\text{HHH, HHT, THT}\}) = \frac{1}{4} = \frac{1}{8} = P(A_1)P(A_2)P(A_3) \).

\( \therefore A_1, A_2, A_3 \) are pairwise independent.

Now, \( P[A_1 \cap A_2 \cap A_3] = P(\{\text{HHH, HHT, THT}\}) \).

\( i = \frac{1}{4} \neq \frac{1}{8} = P(A_1)P(A_2)P(A_3) \).

ii) A: \( \{\text{HHH, THT, THT, TTH}\} \)

B: \( \{\text{HHH, THT, HTH, TTH}\} \)

C: \( \{\text{HHH, HTH, HTT, TTT}\} \)

\( P(A) = P(B) = P(C) = \frac{1}{3} \)

\( P(A \cap B) = \frac{1}{4} = P(A)P(B); P(B \cap C) = \frac{1}{4} = P(B)P(C) \)

\( P(A \cap C) = \frac{1}{4} = P(A)P(C) \).

\( \therefore A, B, C \) are pairwise independent as well as mutual independent.
Important Theorems

1. Define conditional probability. Show that it satisfies all the axioms of probability.

- Conditional Probability

  Classical Def: Conditional probability of the occurrence of the event $B$ given that $A$ has already been occurred, denoted by $P(B|A)$, is defined as,
  
  $$P(B|A) = \frac{N(A \cap B)}{N(A)}$$

  where $N(A)$ is the no. of cases favorable to the event $A$, $N(A \cap B)$ is the no. of cases favorable to the simultaneous occurrence of $A$ and $B$.

  If $N$ be the total no. of equally likely elementary cases then
  
  $$P(B|A) = \frac{N(A \cap B)}{N(A)} = \frac{P(A \cap B)}{P(A)}$$

  \[\Rightarrow P(A \cap B) = P(B|A) P(A)\]

  Axiomatic Def: Consider the probability space $(Ω, Ε, P)$ where $Ω$ is the sample space, $Ε$ is the $σ$-field of the subspace of $Ω$ and $P$ is the probability function defined on $Ε$.

  Let $A \in Ε \Rightarrow P(A) > 0$; then conditional probability of occurrence of any event $B$ belonging to $Ε$ given that $A$ has already been occurred is defined as

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

- Conditional Probability satisfies all the axioms of Probability:

  i. We have $P(A \cap B) \geq 0 \forall B$ and $(B|A)$; and $P(A) > 0$.

  \[\Rightarrow \frac{P(A \cap B)}{P(A)} \geq 0 \forall B\]

  i.e., $P(B|A) \geq 0$ for any $B \in Ε$.

  \[\Rightarrow \text{Axiom I of probability.}\]
ii) Since, \( P(\Omega|\mathcal{A}) = \Omega \),

\[ P(\Omega|\mathcal{A}) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1 \quad (\because P(A) > 0). \]

\[ \Rightarrow \text{Axiom II of probability.} \]

\( \Rightarrow \) Let us consider a sequence of disjoint events \( \{c_n\} \),

\[ c_n \in \mathcal{A} \forall n. \]

Now,

\[ P\left( \bigcup_{n=1}^{\infty} c_n \mid A \right) = \frac{P\left( \bigcup_{n=1}^{\infty} (c_n \cap A) \right)}{P(A)} = \frac{P\left( \bigcup_{n=1}^{\infty} c_n \cap A \right)}{P(A)} = \frac{\sum_{n=1}^{\infty} P(c_n \cap A)}{P(A)} = \sum_{n=1}^{\infty} P(c_n \mid A) \]

\[ \Rightarrow \text{Axiom III of probability.} \]

Hence the proof.

i) What do you mean by stochastic independence of events \( P \)?

\( \Rightarrow \) The event \( A \) is said to be stochastically independent of the event \( B \) if occurrence of \( A \) does not depend upon the occurrence or non-occurrence of \( B \), i.e.,

\[ P(A|B) = P(A) , \quad P(B) > 0. \]

\[ \Rightarrow \frac{P(A \cap B)}{P(B)} = P(A). \]

\[ \Rightarrow P(A \cap B) = P(A)P(B) \quad (i) \]

Similarly \( B \) is said to be stochastically independent of the event \( A \) if,

\[ P(B|A) = P(B) , \quad P(A) > 0. \]

\[ \Rightarrow \frac{P(A \cap B)}{P(A)} = P(B). \]

\[ \Rightarrow P(A \cap B) = P(A)P(B) \quad (ii) \]

Note that the expression (i) is symmetric in \( A \) and \( B \). Hence instead of saying \( A \) is independent of \( B \) on \( B \) is independent of \( A \), one must say \( A \) and \( B \) are independent of each other.

Remark: If two events are mutually exclusive then they will not be stochastically independent of each other.
3) State and prove Compound Probability Theorem.

**Statement:** The probability of simultaneous occurrence of A and B is given by the product of the unconditional probability of the event A by the conditional probability of B, supposing that A actually occurred. In other words,

\[ P(A \cap B) = P(A)P(B|A). \]

**Proof:**

Let there be \( n \) no. of all possible outcomes, of these,

- \( n_A \) = no. of outcomes favorable to A.
- \( n_B \) = no. of outcomes favorable to B.
- \( n_{AB} \) = no. of outcomes favorable to A and B.

Then, \( P(A) = \frac{n_A}{n} \), \( P(A \cap B) = \frac{n_{AB}}{n} \) and \( P(B|A) = \frac{n_{AB}}{n_A} \).

\[ P(A \cap B) = \frac{n_{AB}}{n} = \frac{n_A}{n} \times \frac{n_{AB}}{n_A} \text{ [It is supposed that A has actually been occurred, i.e., } P(A) > 0 \text{ and hence } n_A > 0] \]

Hence the theorem is proved.

In general case, if \( A_1, \ldots, A_n \) be any events in \( \mathcal{F} \), then by induction

\[ P\left( \bigcap_{i=1}^{n} A_i \right) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap \cdots \cap A_{n-1}), \]

provided \( P(A_1 \cap \cdots \cap A_{n-1}) > 0 \).

This is called 'Law of Multiplication'.
4) State & proof the theorem of total probability.

Set $\Omega \rightarrow$

Total Probability

**Theorem:** Let $(\Omega, \mathcal{E}, P)$ be the probability space, suppose $\{H_n\}$ is a sequence of mutually exclusive and exhaustive events such that $P(H_n) > 0 \forall n$, $H_n \in \mathcal{E} \forall n$.

Then the probability of any event $B \in \mathcal{E}$ is given by

$$P(B) = \sum_{n=1}^{\infty} P(H_n)P(B \mid H_n)$$

**Proof:** Since $\{H_n\}$ is a sequence of mutually exclusive and exhaustive events,

$$\bigcup_{n=1}^{\infty} H_n = \Omega$$

Now, $B = B \cap \Omega$.

$$\therefore P(B) = P\left[B \cap \left(\bigcup_{n=1}^{\infty} H_n\right)\right] = P\left[\bigcup_{n=1}^{\infty} (B \cap H_n)\right]$$

Note that $H_i \cap H_j = \emptyset \forall i \neq j$

$$\therefore (B \cap H_i) \cap (B \cap H_j) = \emptyset \forall i \neq j.$$ Clearly, $\{B \cap H_n\}$ is also a sequence of mutually disjoint events $\in \mathcal{E}$. Hence by Axiom-III, we have

$$P\left(\bigcup_{n=1}^{\infty} (B \cap H_n)\right) = \sum_{n} P(B \cap H_n)$$

Thus,

$$P(B) = \sum_{n} P(B \cap H_n)$$

So,

$$P(B) = \sum_{n} P(H_n)P(B \mid H_n) \quad [\text{From the axiom of compound probability}]$$

Hence the proof.

**Implication:** The implication of this result is that the unconditional probability of the event $B$ can be obtained as the weighted average of the conditional probabilities.
Application of Total Probability Theorem:

1. A box has 12 red and 6 black balls. A ball is selected from the box. If it is red, it is returned to the box. If the ball is black, it and 2 additional balls are added to the box. Find the probability that a second ball drawn from the box is (i) red (ii) black.

Sol. Let Ri and Bi respectively be the event that the ith ball drawn is red and that the ith ball drawn is black for i = 1, 2.

\[ P(R_1) = \frac{12}{18}, \quad P(B_1) = \frac{6}{18} \]
\[ P(R_2 | R_1) = \frac{12}{18}, \quad P(R_2 | B_1) = \frac{12}{20} \]
\[ P(B_2 | R_1) = \frac{6}{18}, \quad P(B_2 | B_1) = \frac{8}{20} \]

(i) \[ P(R_2) = P(R_1) P(R_2 | R_1) + P(B_1) P(R_2 | B_1) \]
\[ = \frac{12}{18} \times \frac{12}{18} + \frac{6}{18} \times \frac{12}{20} \]
\[ = \frac{29}{45} \]

(ii) \[ P(B_2) = P(R_1) P(B_2 | R_1) + P(B_1) P(B_2 | B_1) \]
\[ = \frac{12}{18} \times \frac{6}{18} + \frac{6}{18} \times \frac{8}{20} \]
\[ = \frac{16}{45} \]

2. Let the probability \( pn \) that a family has \( n \) children be \( \alpha^n \) when \( n \geq 1 \) and let \( p_0 = 1 - \alpha (1 + p + p^2 + \ldots) \). Suppose that a child is as likely to be a male as to be a female. Show that for \( k \geq 1 \) the prob. that a family contains exactly \( k \) boys is \( \frac{2\alpha^k}{(2-p)^{k+1}} \).

Sol. Let \( B_n \) denote the event that the family contains \( n \) children and \( A_k \) denote the event that it has \( k \) boys. The probability we require is \( P(A_k) \).
Note that $B_n (n = 0, 1, 2, \ldots)$ are exhaustive as well as mutually exclusive. So we apply the theorem of total probability to get

$$P(A_k) = \sum_{n=0}^{\infty} P(B_n) P(A_k \mid B_n)$$

Given that $P(B_n) = \alpha p^n$, where $P_0 = 1 - \alpha p (1 + p + p^2 + \ldots)$

$$P(A_k \mid B_n) = P(k \text{ boys} \mid n \text{ children}) = \begin{cases} \binom{n}{k} \left(\frac{1}{2}\right)^n & \text{for } n \geq k \\ 0 & \text{for } n < k \end{cases}$$

Hence by Total Probability theorem,

$$P(A_k) = \sum_{n=0}^{\infty} P(B_n) P(A_k \mid B_n)$$

$$= \sum_{n=0}^{\infty} P(B_n) P(A_k \mid B_n)$$

$$= \sum_{n=k}^{\infty} \alpha p^n \binom{n}{k} \left(\frac{1}{2}\right)^n$$

$$= \alpha \left(\frac{p}{2}\right)^k \sum_{n=k}^{\infty} \binom{n}{k} \left(\frac{p}{2}\right)^{n-k}$$

$$= \alpha \left(\frac{p}{2}\right)^k \left(1 - \frac{p}{2}\right)^{-(k+1)}$$

$$= \frac{2\alpha p^k}{(2 - \phi)^{k+1}}.$$
5. State and prove Bayes' theorem.

**Statement:** (Bayes' Theorem)

For a sequence of mutually exclusive and exhaustive events \( A_1, A_2, \ldots \in \Omega \) with \( P(A_i) > 0 \) for all \( i \),

\[
P(A_j \mid B) = \frac{P(A_j) P(B \mid A_j)}{\sum_{i=1}^{\infty} P(A_i) P(B \mid A_i)}, \quad \text{where } B \text{ is any other event.}
\]

**Proof:** Since \( A_1, A_2, \ldots \) are mutually exclusive and exhaustive events, \( P(A_i) > 0 \),

\[
\sum_{i=1}^{\infty} P(A_i) = P\left( \bigcup_{i=1}^{\infty} A_i \right) = P(\Omega) = 1.
\]

Let \( B = B \cap \Omega = B \cap \left( \bigcup_{i=1}^{\infty} A_i \right) = \bigcup_{i=1}^{\infty} (B \cap A_i) \).

\[
\therefore P(B) = P\left( \bigcup_{i=1}^{\infty} (B \cap A_i) \right) \quad \text{[(B \cap A_i) is a sequence of mutually disjoint events in } \Omega, \text{ applying Axiom III]}.
\]

\[
= \sum_{i=1}^{\infty} P(B \cap A_i) = \sum_{i=1}^{\infty} P(A_i) P(B \mid A_i)
\]

\[
= \sum_{i=1}^{\infty} P(A_i) P(B \mid A_i) \quad (1)
\]

Now, \( P(A_j \mid B) = \frac{P(A_j \cap B)}{P(B)} \)

\[
= \frac{P(A_j) P(B \mid A_j)}{\sum_{i=1}^{\infty} P(A_i) P(B \mid A_i)}, \quad \text{by (1) and since } P(B) > 0.
\]

Hence the theorem is proved.
Application of Baye’s Theorem:

(1) In answering a question on a multiple-choice test, an examinee either knows the answer (with probability \( p \)) or guesses (with probability \( 1-p \)). Let the probability of guessing be \( q \) for one who guesses the answer correctly and \( 1/m \) for one who knows the answer. Now suppose an examinee answers a question correctly. What’s the prob. that he really knows the answer?

**Solution:** Let \( A_1 \) be the event that an examinee knows the answer, \( A_2 \) be the event that the answer is correct. Then
\[
P(A_1) = p, \quad P(A_2) = 1-p, \quad P(B|A_1) = 1, \quad P(B|A_2) = \frac{1}{m}.
\]
By Baye’s theorem,
\[
P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}
\]
\[
= \frac{p \cdot 1}{p \cdot 1 + (1-p) \cdot \frac{1}{m}}
\]
\[
= \frac{mp}{1+(m-1)p}
\]

(2) There are also two drawers in each of three boxes that are identical in appearance. The first one contains a gold coin, the second contains a silver coin in each drawer, but the third contains a gold coin in one drawer and a silver coin in the other. A box is chosen, one of its drawers is opened and a gold coin is found. What’s the probability that the other drawer too will have a gold coin?

**Solution:** Let \( A_1 \) be the event that the first box is chosen, \( A_2 \) be the event that the second box is chosen, \( A_3 \) be the event that the third box is chosen, and \( B \) be the event that the second one is a gold coin.
\[
P(B|A_1) = 1, \quad P(B|A_2) = 0, \quad P(B|A_3) = \frac{1}{2}.
\]
By Baye’s theorem,
\[
P(A_1|B) = \sum_{i=1}^{3} \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{3} P(A_i)P(B|A_i)}
\]
\[
= \frac{\frac{1}{3} \times 1}{\frac{1}{3} \times 1 + \frac{1}{3} \times 0 + \frac{1}{3} \times \frac{1}{2}}
\]
\[
= \frac{2}{3}.
\]
In a doll factory, machines $M_1$, $M_2$, and $M_3$ manufacture respectively 45, 25, and 30 percent of the total output. Of their output, 6, 8, and 3 percent respectively are defective. What is the probability that it was manufactured by $M_1$?

**Sol.** Bi: Chosen ball is manufactured by machine $M_i$;
A: Chosen doll is defective.

By Baye's Theorem,

$$P(B_i | A) = \frac{P(B_i) P(A | B_i)}{\sum_{i=1}^{3} P(B_i) P(A | B_i)}, \quad i = 1, 2, 3$$

$$P(M_1 | A) = \frac{45 \times 6}{45 \times 6 + 25 \times 8 + 30 \times 3} = \frac{27}{54}.$$

An urn containing 5 balls has been filled up by taking 5 balls from another urn containing 5W & 5B balls.

A ball is taken at random from urn 1 and it happens to be black. What's the prob. of drawing a cohtine ball from the remaining?

**Sol.** Let Bi denotes that among 5 balls kept in urn 1, exactly $i$ are white.

A: The first ball taken from urn 1 is black,
C: The second ball drawn from urn 1 is white.

$$P(C | A) = \sum_{i=0}^{5} \frac{P(B_i) P(A | B_i) P(C | A \cap B_i)}{P(B_i) P(A | B_i)}$$

$$P(B_i) = \binom{5}{i} \binom{5}{5-i} / \binom{10}{5}, \quad i = 0(1) 5$$

$$P(A | B_i) = \frac{5-i}{5}, \quad P(C | A \cap B_i) = \frac{4}{4}, \quad i = 0(1) 4.$$

| i  | $P(B_i)$ | $P(A | B_i)$ | $P(C | A \cap B_i)$ |
|----|----------|-------------|---------------------|
| 0  | 1/10     | 1           | 0                   |
| 1  | 25/100   | 1/5         | 1/4                 |
| 2  | 100/1000 | 3/5         | 1/2                 |
| 3  | 100/1000 | 2/5         | 3/4                 |
| 4  | 25/1000  | 1/5         | 1                   |
| 5  | 1/10     | 0           |                     |

$$\therefore P(C | A) = \frac{5/18}{1/2} = \frac{5}{9}.$$
6) What is the probability of getting 's' points when
n die is rolled.

\[ \text{Sol.} \Rightarrow \]
Here, no. of all possible cases is \(6^n\).
Since, the dice are fair, so all possible cases are equally likely.
The no. of favorable cases of 'getting a sum s' = The no. of solutions \((n_1, n_2, \ldots, n_k)\) of the equation
\[ \sum n_i = s, \quad n_i \geq 0 \]
= The coefficient of \(x^s\) in the expansion of
\[ (x^1 + x^2 + \ldots + x^6)^n \]
= The coefficient of \(x^s\) in \(\frac{x^s(1-x^6)}{1-x}\)
= The coefficient of \(x^{s-n}\) in the expansion of
\[ (1-x^6)^n \]
= \(\binom{s-n}{s-n}\)
= The coefficient of \(x^s\) in \(\sum_{i=0}^{\left\lfloor \frac{s-n}{6} \right\rfloor} \binom{s-n}{i} (-1)^i (s-n-i) \)
= \(\sum_{i=0}^{\left\lfloor \frac{s-n}{6} \right\rfloor} (-1)^i (s-n-6i) \)
= The required probability is \(\frac{\sum_{i=0}^{\left\lfloor \frac{s-n}{6} \right\rfloor} (-1)^i (s-n-6i)}{6^n}\)

Cor. Show that the prob. of getting a total of s with n die is same as the prob. of throwing \(7(n-s)\).

\[ \text{Sol.} \]
The co-efficient of \(x^s\) in \((x + x^2 + \ldots + x^6)^n\)
= \(\binom{s}{1} (1-x^6)^n \)
= \(x^s \binom{s}{s-n} \)
= \(x^{s-n} \)
= \(\binom{s-n}{s-n} \)

\[(x + x^2 + \ldots + x^6)^n\]
State & prove Boole's Inequality.

Statement: (Boole's Inequality) If $A_i$ (i=1(1)n),
be any events in $\mathbb{A}$, then

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

Proof: Consider first $A_1$ and $A_2$. Now

$$A_1 \cup A_2 = A_1 \cup (A_2 - A_1),$$

the events $A_1$, $(A_2 - A_1)$ being disjoint. Hence

$$P(A_1 \cup A_2) = P(A_1) + P(A_2 - A_1) \leq P(A_1) + P(A_2).$$

Since $(A_2 - A_1) \subseteq A_2$, thus the inequality is proved for $n=2$. Now

$$\bigcup_{i=1}^{n} A_i = \left(\bigcup_{i=1}^{n-1} A_i\right) \cup A_n$$

$$\Rightarrow P\left(\bigcup_{i=1}^{n} A_i\right) \leq P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n) \leq P\left(\bigcup_{i=1}^{n-1} A_i\right) + P(A_n - A_{n-1}) + P(A_n).$$

Hence

$$\leq \sum_{i=1}^{n} P(A_i)$$

"=" holds iff $A_i \cap A_j = \emptyset \forall i, j = 1(1)n$. 
State and proof Bonferroni's inequality.

Statement: (Bonferroni's Inequality)

If \( A_i \) for \( i = 1 \ldots K \) be events in \( \mathcal{A} \), then
\[
P \left( \bigcap_{i=1}^{K} A_i \right) \geq \sum_{i=1}^{K} P(A_i) - (K-1).
\]

Proof: \( \Rightarrow \) We have \( \left( \bigcap_{i=1}^{K} A_i \right) = \left( \bigcup_{i=1}^{K} A_i^c \right)^c \)
\[= P \left( \bigcap_{i=1}^{K} A_i \right) = 1 - P \left( \bigcup_{i=1}^{K} A_i^c \right) \quad 0 \]

By Boole's inequality,
\[P \left( \bigcup_{i=1}^{K} A_i^c \right) \leq \sum_{i=1}^{K} P(A_i^c) = \sum_{i=1}^{K} \left[ 1 - P(A_i) \right] = K - \sum_{i=1}^{K} P(A_i).\]

\[\Rightarrow 1 - P \left( \bigcap_{i=1}^{K} A_i \right) \leq K - \sum_{i=1}^{K} P(A_i) \quad [\text{Applying 0}].\]

\[\Rightarrow P \left( \bigcap_{i=1}^{K} A_i \right) \geq \sum_{i=1}^{K} P(A_i) - (K-1)\]

\( \Rightarrow \) holds iff \( A_i \cap A_j = \emptyset \ \forall i \neq j = 1(1)K \).

i.e. if chosen events from \( \mathcal{A} \) are mutually exclusive.
Statement: - (Poincaré's Theorem) For any \( n \geq 2 \) events 
\( A_1, A_2, \ldots, A_n \), not necessarily mutually exclusive,
the probability of occurrence of at least one of these events will be given by:
\[
P\left( \bigcup_{i=1}^{n} A_i \right) = \sum_{i=1}^{n} P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \ldots + (-1)^{n-1} P(A_1 \cap \cdots \cap A_n)
\]

Proof: \( \rightarrow \) First consider two events \( A_1 \) and \( A_2 \).
Since \( A_1 \cup A_2 = A_1 + (A_2 - A_1 \cap A_2) \]
\[
P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)
\]
Since \( A_1 \cap A_2 \subseteq A_2 \)
\[
P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)
\]
The result is true for \( n = 2 \).
Now, assumed that the result is true for \( n = m \)
So, \( P\left( \bigcup_{i=1}^{m} A_i \right) = \sum_{i=1}^{m} P(A_i) - \sum_{1 \leq i < j \leq m} P(A_i \cap A_j) + \ldots + (-1)^{m-1} P(A_1 \cap \cdots \cap A_m) \)

Now, let us include one more event \( A_{m+1} \) in \( \mathcal{A} \).
Now, \( \bigcup_{i=1}^{m+1} A_i = (\bigcup_{i=1}^{m} A_i) \cup A_{m+1} \)
\[
P\left( \bigcup_{i=1}^{m+1} A_i \right) = P\left( \bigcup_{i=1}^{m} A_i \right) + P(A_{m+1}) - P\left( (\bigcup_{i=1}^{m} A_i) \cap A_{m+1} \right)
\]
\[
= P\left( \bigcup_{i=1}^{m} A_i \right) + P(A_{m+1}) - P\left( \bigcup_{i=1}^{m} (A_i \cap A_{m+1}) \right)
\]
\[
= \left[ \sum_{i=1}^{m} P(A_i) - \sum_{1 \leq i < j \leq m} P(A_i \cap A_j) + \ldots + (-1)^{m-1} P\left( \bigcap_{i=1}^{m} A_i \right) \right]
\]
\[
+ P(A_{m+1}) - \left[ \sum_{i=1}^{m} P(A_i \cap A_{m+1}) - \sum_{1 \leq i < j \leq m} P(A_i \cap A_j \cap A_{m+1}) + \ldots + (-1)^{m-1} P\left( \bigcap_{i=1}^{m+1} A_i \right) \right]
\]
\[
= \sum_{i=1}^{m+1} P(A_i) - \sum_{1 \leq i < j \leq m+1} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq m+1} P(A_i \cap A_j \cap A_{k+m+1}) + \ldots + (-1)^{m} P\left( \bigcap_{i=1}^{m+1} A_i \right)
\]
\[
\therefore \text{The theorem is true for } n = m+1, \text{ when it is true for } n = m. \text{ Hence by induction, result follows.}
Proof: First consider any sequence of non-decreasing events. Then $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$.

\[ \lim_{n \to \infty} A_n = \bigcup_{n=1}^{\infty} A_n. \]

\[ = \bigcup_{n=1}^{\infty} \left( A_n - \bigcup_{i=1}^{n-1} A_i \right). \]

\[ = \bigcup_{n=1}^{\infty} (A_n - A_{n-1}). \]

Since $A_i$'s are non-decreasing, so

\[ \bigcup_{i=1}^{n-1} A_i = A_{n-1}. \]

\[ \therefore P(\lim_{n \to \infty} A_n) = P \left[ \bigcup_{n=1}^{\infty} (A_n - A_{n-1}) \right] \]

\[ = \bigcup_{n=1}^{\infty} P(A_n - A_{n-1}) \]

[as $(A_n - A_{n-1})$ are disjoint events, so by the axiom of countable additivity of $P(\cdot)$].

\[ = \sum_{n=1}^{\infty} \left[ P(A_n) - P(A_{n-1}) \right] \]

\[ = \lim_{n \to \infty} \sum_{n=1}^{\infty} \left[ P(A_n) - P(A_{n-1}) \right] \]

[\( \therefore A_{n-1} \subseteq A_n \), so therefore $P(\cdot)$ is subtrractive]

\[ = \lim_{n \to \infty} P(A_n). \]

Now consider any sequence of non-decreasing increasing events. Then, $A_n$ will be monotone non-decreasing and each will belong to $\mathfrak{E}$.

\[ \therefore P\left( \lim_{n \to \infty} A_n \right) = \lim_{n \to \infty} P\left( A_n \right) \]

\[ \Rightarrow P\left( \lim_{n \to \infty} A_n \right)^c = \lim_{n \to \infty} (1 - P\left( A_n \right)). \]

\[ \Rightarrow 1 - P\left( \lim_{n \to \infty} A_n \right) = 1 - \lim_{n \to \infty} P\left( A_n \right). \]

i.e., $P\left( \lim_{n \to \infty} A_n \right) = \lim_{n \to \infty} P\left( A_n \right).$

\[ \therefore \] Hence the result is proved.

**Statement (Continuity theorem)**

If $\{A_n\}$ is a monotone sequence of events belonging to $\mathfrak{E}$, then

\[ P\left( \lim_{n \to \infty} A_n \right) = \lim_{n \to \infty} P\left( A_n \right). \]
Derive an expression for the probability of realization of exactly 'm' out of 'n' events.
(Jordon's Theorem).

Statement: Probability that exactly m of the events $A_i$, $i = 1(1)n$, will occur is

$$P[m] = s_m - \binom{m+1}{m} s_{m+1} + \binom{m+2}{m} s_{m+2} - \cdots$$

where $s_1 = \sum_{i=1}^{n} P(A_i)$, $s_2 = \sum_{i<j}^{n} P(A_i \cap A_j)$ and so on.

Proof: Consider first the probability that just m specified events among $A_i$, $i = 1(1)n$, will occur, say the events $A_1, A_2, \ldots, A_m$. The probability is

$$P(A_1 \cap A_2 \cap \ldots \cap A_m \cap A_1^c \cap A_2^c \cap \ldots \cap A_n^c)$$

Now, take $A_1 \cap A_2 \cap \ldots \cap A_m \cap B$.

The above equals to

$$P[B \cap (\bigcup_{i=m+1}^{n} A_i^c)]$$

$$= P(B) - P[B \cap (\bigcup_{i=m+1}^{n} A_i^c)]$$

$$= P(B) - P[\bigcup_{i=m+1}^{n} (B \cap A_i^c)]$$

$$= P(\bigcup_{i=1}^{m} A_i) - \sum_{i} P(B \cap A_i) + \sum_{i<j} P(B \cap A_i \cap A_j)$$

$$\cdots - \cdots + (-1)^{n-m} P(\bigcap_{i=1}^{m} A_i^c) \quad \text{(*)}$$

We may choose m specified events out of 'n' events in $\binom{n}{m}$ mutually exclusive ways. So, the required probability is the sum of $\binom{n}{m}$ such terms. Again, each such probability has $\binom{n-m}{m}$ terms included in $s_{m+i}$ and sign attached to it is $(-1)^i$ while the total number of terms in $s_{m+i}$ is $\binom{n}{m+i}$. Hence, the coefficient of $s_{m+i}$ in the expression (*) is
\[ (-1)^i \frac{(\begin{array}{c} n \\ m \end{array}) (\begin{array}{c} n-m \\ r \end{array})}{(\begin{array}{c} n \\ m-i \end{array})} \]

Therefore,
\[ P_{[m]} = S_m - \frac{(\begin{array}{c} m+1 \\ m \end{array})}{(\begin{array}{c} m+1 \\ m \end{array})} S_{m+1} + \frac{(\begin{array}{c} m+2 \\ m \end{array})}{(\begin{array}{c} m+2 \\ m \end{array})} S_{m+2} - \cdots \]
\[ + (-1)^{n-m} \frac{(\begin{array}{c} n \\ m \end{array})}{(\begin{array}{c} n \\ m \end{array})} S_n. \]

**Theorems**

**Problem 1**

Discuss the method of determination of the probability that \( m (\geq 2) \) or more of the events occur simultaneously out of events \( A_1, A_2, \ldots, A_n \).

Illustrate with an example.

**Solution:** At least \( m \) events occur if and only if \( m+1 \) events occur, \( i = 0, 1, 2, \ldots, n-m \).

\[ \therefore P_m = P_{[m]} + P_{[m+1]} + \cdots + P_{[n]} = S_m - \sum \frac{\begin{array}{c} m+1 \\ m \end{array}}{\begin{array}{c} m+1 \\ m \end{array}} S_{m+1} + \cdots \]
\[ + (-1)^{n-m} \frac{\begin{array}{c} n \\ m \end{array}}{\begin{array}{c} n \\ m \end{array}} S_n. \]

[Using Jordan's first theorem].

Coefficient of \( S_{m+i} \) in the R.H.S. is
\[ = (-1)^i \sum_{j=0}^i (-1)^j \frac{\begin{array}{c} m+j \\ m+i \end{array}}{\begin{array}{c} m+j \\ m+i \end{array}}. \]

\[ = \sum_{j=0}^i (-1)^j \left( \begin{array}{c} m+i-j \\ m+j \end{array} \right) + \begin{array}{c} m+i-j \\ m-j \end{array}. \]

\[ = \left( \begin{array}{c} m+i-1 \\ m \end{array} \right) + \begin{array}{c} m+i-1 \\ m-1 \end{array} - \left( \begin{array}{c} m+i-1 \\ m \end{array} \right) \cdot \begin{array}{c} m+i-1 \\ m-1 \end{array} \]
\[ + \left( \begin{array}{c} m+i-1 \\ m+2 \end{array} \right) + \begin{array}{c} m+i-1 \\ m+1 \end{array} \cdot \begin{array}{c} m+i-1 \\ m+1 \end{array} + \cdots + (-1)^i \begin{array}{c} m+i-1 \\ m+i-1 \end{array} \]
\[ = \begin{array}{c} m+i-1 \\ m-1 \end{array} \cdot \begin{array}{c} m+i-1 \\ m-1 \end{array} \cdot \begin{array}{c} m+i-1 \\ m+1 \end{array} \cdot \begin{array}{c} m+i-1 \\ m+1 \end{array} \cdot \cdots \cdot (-1)^i \begin{array}{c} m+i-1 \\ m+i-1 \end{array} \cdot \begin{array}{c} m+i-1 \\ m+i-1 \end{array} \]
**SELECTED PROBLEMS:** (Application of Jordan's Theorem)

1. **(Matching Problem):** n letters are placed at random into similarly addressed n envelopes. Find the probability of exactly m matches. Also

   **Solution:** Jordan's Theorem states,

   \[ P(\text{occurrence of exactly } m \text{ of } n \text{ events}) = P_{m} = \sum_{i=0}^{n-m} (-1)^{i} \binom{m+i}{m} \times S_{m+i} \]

   \[ S_{k} = \sum_{1 \leq i_{1} < i_{2} < \cdots < i_{k} \leq n} P(A_{i_{1}}, A_{i_{2}} \ldots, A_{i_{k}}) \]

   Note: \[ P(A_{i}) = \frac{(n-1)!}{n!} = \frac{1}{n} \quad \forall \ i \]

   \[ S_{1} = \sum_{i=1}^{n} P(A_{i}) = n \cdot \frac{(n-1)!}{n!} \]

   \[ P(A_{i} \cap A_{j}) = \frac{(n-2)!}{n!} \quad \forall \ i \neq j \]

   \[ S_{2} = \sum_{1 \leq i < j \leq n} P(A_{i} \cap A_{j}) = \binom{n}{2} \cdot \frac{(n-2)!}{n!} \]

   \[ \Rightarrow S_{k} = \binom{n}{k} \frac{(n-k)!}{n!} = \frac{1}{k!} \]

   \[ \Rightarrow S_{m+i} = \frac{(m+i)!}{(m+i)!} \]

   \[ \therefore P(\text{exactly } m \text{ matches}) = \sum_{i=0}^{n-m} (-1)^{i} \binom{m+i}{m} \frac{(n-m)!}{i!} \]

   \[ = \frac{1}{m!} \sum_{i=0}^{n-m} \frac{(-1)^{i}}{i!} \]

   \[ = e^{-1} \quad \text{, since } n \uparrow \infty \]

   \[ \therefore P(\text{at least } m \text{ matches}) = \sum_{i=0}^{n-m} (-1)^{i} \binom{m+i-1}{m-1} S_{m+i} \]

   \[ = \sum_{i=0}^{n-m} (-1)^{i} \binom{m+i-1}{m-1} \frac{1}{(m+i)!} \]

   \[ = \frac{1}{(m-1)!} \sum_{i=0}^{n-m} (-1)^{i} \frac{1}{(m+i)! \cdot i!} \]
SOME SOLVED EXAMPLES:

Q.1. Two players A and B have respectively \((n+1)\) and \(n\) coins. If they toss their coins simultaneously, what is the probability that
(a) A will have more heads than B
(b) A will have as many heads as B
(c) A will have fewer heads than B

Solution:

\(X\) : Number of heads obtained by A
\(Y\) : Number of heads obtained by B

Assuming the coin is fair, so

\[
P(A \text{ will have } x \text{ heads and } B \text{ will have } y \text{ heads}) = \binom{n+1}{x} \left(\frac{1}{2}\right)^{n+1} \binom{n}{y} \left(\frac{1}{2}\right)^n = \frac{1}{2^{2n+1}} \binom{n+1}{x} \binom{n}{y}.
\]

(a) \(P(X > Y) = \sum_{x > y} \frac{1}{2^{2n+1}} \binom{n+1}{x} \binom{n}{y}\)

(b) \(P(X = Y) = \sum_{x \leq y} \frac{1}{2^{2n+1}} \binom{n+1}{x} \binom{n}{y}\)

\(= \frac{1}{2^{2n+1}} \sum_{x' > y} \binom{n+1}{x'} \binom{n}{y}\)

\(= P(X > Y) = \frac{1}{2}\).

(c) \(P(X < Y) = 1 - P(X > Y) = 1 - P(X > Y) - P(X = Y) = 1 - \frac{1}{2} - \frac{1}{2^{2n+1}} \binom{2n+1}{n}.
\)
Q.2. In a game called 'odd men out', n persons toss coin to determine one person who will buy refreshments for the whole group. If there is a person in the group whose outcome (head/tail) is different from that of any other member in the group, then that person is an odd man. What's the probability that —

(a) In a game there will be an odd man.
(b) n plays will be required to conclude the game.

Solution:

The odd man will be found if

(a) i. all the other (n-1) members will get a Head and the remaining one gets tail and
ii. other (n-1) members will get a Tail and the remaining will get a Head.

\[ P(\text{i-th part}) = \binom{n}{n-1} \left(\frac{1}{2}\right)^n \]

\[ P(\text{ii-th part}) = \binom{n}{n-1} \left(\frac{1}{2}\right)^n \]

\[ P(\text{there will be an odd man in a game}) = 2 \left(\binom{n}{n-1} \left(\frac{1}{2}\right)^n \right) \]

\[ = \frac{n}{2^{n-1}} \]

(b) p plays will be required if the first (n-1) plays will not give an odd man and the nth one will give an odd man. Since the plays are independent,

Required probability = \[ \int_{1-p}^{1} \frac{1}{2^{n-1}} \left[ \frac{n}{2^{n-1}} \right] \]
Q. 3. A fair die is thrown 7 times. What is the probability of getting a total of 30 points.

Solution: Each throw of the die will give any of the numbers 1, 2, 3, ..., 6.

Therefore, the total number of elementary events in 7 throws is $6^7$. The die is fair implies that each of these $6^7$ elementary events will be equally likely.

Let $A$ denote the event that the sum of the points is 30. Then the number of elementary events favourable to $A$ is the number of solutions of $x_1 + x_2 + \ldots + x_7 = 30$.

The number of solutions is the same as the coefficient of $t^{30}$ in the expansion of

$$
(1 + t + t^2 + \ldots + t^6)^7 = t^7 \left(\frac{1 + t + t^2 + \ldots + t^5}{1 - t}\right)^7 = t^7 \left(\frac{1 - t^6}{1 - t}\right)^7, \quad |t| < 1
$$

This is the same as the coefficient of $t^{23}$ in

$$
(1 - 7t^6 + \frac{7}{2}t^{12} - \frac{7}{3}t^{18} - \ldots - t^{42})(1 + 7t^3 + \frac{7}{2}t^6 + \frac{7}{3}t^9 + \ldots)
$$

This is again equal to

$$
\binom{7.8.9 \ldots 29}{23!} - 7\left(\binom{7.8.9 \ldots 23}{17!}\right) + \frac{7}{2} \left(\binom{7.8.9 \ldots 17}{11!}\right) - \frac{7}{3} \left(\binom{7.8.9 \ldots 11}{5!}\right) = n_1
$$

So, required prob. $= \frac{n_1}{6^7}$
Q.4. A man addresses in \( n \) envelops and comite \( n \) cheques in payment of \( n \) bills. If the \( n \) bills are placed at random in the \( n \) envelopes, what is the prob. that each bill coill be placed in a corong envelopes.

**Solution:**

\( A_i \) : The event that \( i^{th} \) bill goes to the \( i^{th} \) envelope is

Then required probability is 

\[
P(\bigcap_{i=1}^{n} A_i^c) = P(\bigcup_{i=1}^{n} A_i)
\]

\[
= 1 - P(\bigcap_{i=1}^{n} A_i)
\]

\[
= 1 - S_1 + S_2 - S_3 + \cdots + (-1)^n S_n
\]

where 

\[
S_i = \sum_{i=1}^{n} P(A_i)
\]

\[
S_2 = \sum_{i<j} P(A_i \cap A_j)
\]

\[
S_n = P(A_1 \cap A_2 \cdots \cap A_n)
\]

Now \( n \) bills can be placed in \( n \) envelopes in \( n! \) ways. These placements are made at random means that these \( n! \) arrangements are all equally likely. Now in order that \( A_i \) has to occur, the \( i^{th} \) bill will go to the \( i^{th} \) envelope and the remaining \((n - 1)\) bills can be placed among the \((n - 1)\) envelopes in \((n - 1)!\) ways.

\[
P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}
\]

By a similar argument, 

\[
P(A_i \cap A_j) = \frac{1}{n(n-1)} \text{ for } i < j
\]

\[
P(A_i \cap A_j \cap A_k) = \frac{1}{n(n-1)(n-2)} \text{ for } i < j < k
\]

\[
\therefore \text{ Required probability is equal to }
\]

\[
1 - \left(\frac{n}{1}\right) \frac{1}{n} + \left(\frac{n}{2}\right) \frac{1}{n(n-1)} - \cdots + (-1)^n \frac{1}{n!}
\]

\[
= 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n \frac{1}{n!}
\]

\[
\to e^{-1} \text{ when } n \to \infty
\]
Ex. 5: Three prisoners, one of whom may call A, B and C, are informed by the jailer that one of them has been chosen at random to be executed and the other two are to be freed. Prisoner A who has studied probability theory, then reasons to himself that he has probability 1/3 of being executed. He then asks the jailer to tell him privately which of his fellow prisoners will be free, claiming that there would not be any harm in divulging this information, since he already knows that at least one will go. The jailer being an ethical fellow refuses to reply to the question, pointing out that if A knew which of his fellows were to be set free, then his prob. of being executed would increase to 2/3, since he would then be executed. One of two prisoners, one of whom is to be executed, is still 1/3, even if the jailer were to answer his question, 1/3, assuming that in the event that A is to be executed, the jailer is as likely to say that B is to be set free as he is to say that C is to be set free.

Solution: Let $E_i$: the event that prisoner $i$ will be executed, for $i = A, B, C$.

Let $F_i$: the event that prisoner $i$ will be set free, for $i = A, B, C$.

Compute,

1. $P(E_A) = 1/3$ since if the decision is to execute A then the jailer is as likely to say B will be set free as to say C will be set free.

2. $P(F_B | E_A) = 1/2$ since the jailer can't tell a lie.

3. $P(F_B | E_C) = 1$ since the jailer can't tell A that he will be set free.

By Bayes' theorem,

$$P(E_A | F_B) = \frac{P(E_A) P(F_B | E_A)}{P(E_A) P(F_B | E_A) + P(E_B) P(F_B | E_B) + P(E_C) P(F_B | E_C)}$$

$$= \frac{1/3 \times 1/2}{1/3 \times 1/2 + 0 \times 1/3 + 1 \times 1/3} = 1/3.$$
Ex. 6. There are $N$ coupons numbered $1, 2, \ldots, N$ in a box. If $n$ coupons are drawn at random, then what's the probability that the highest number on coupon drawn is $m$.

Ans. $A_i$: the event that all coupons drawn have numbers not exceeding $i$.

**Case I:** (Drawing WR)

Required prob. $= P(A_m) - P(A_{m-1})$

$= \frac{m^n - (m-1)^n}{N^n}$

**Case II:** Required prob. $= P(A_m) - P(A_{m-1})$

$= \frac{(m)^n - (m-1)^n}{N^n}$

Ex. 7. A fair dice is thrown $n$ times. What's the probability that each of the six numbers $1, 2, 3, \ldots, 6$ will appear at least once.

Solution: Let $A_i$ denote the event that the $i$th variety will appear at least once, $i=1(0)6$.

Then the probability we require is $P\left(\bigcap_{i=1}^{6} A_i\right)$

$= P\left(\bigcup_{i=1}^{6} A_i^c\right)^c$

$= 1 - P\left(\bigcup_{i=1}^{6} A_i^c\right)$

$= 1 - \sum_{i=1}^{6} P(A_i^c) - \sum_{i<j}^{6} P(A_i^c \cap A_j^c) + \ldots + \sum_{i_1<i_2<\ldots<i_6} P(A_{i_1}^c \cap A_{i_2}^c \cap \ldots \cap A_{i_6}^c)$

[By Poincare's Theorem]

In a single throw of dice, prob. that the number $1$ will not appear is $\frac{5}{6}$. $P(A_i^c) = \left(\frac{5}{6}\right)^n$, since throws are independent.

Similarly, $P(A_i^c \cap A_j^c) = \left(\frac{4}{6}\right)^n$.

$P(A_{i_1}^c \cap A_{i_2}^c \cap \ldots \cap A_{i_6}^c) = \left(\frac{1}{6}\right)^n$

So, required prob. is

$= 1 - \left(\frac{6}{1}\right)\left(\frac{5}{6}\right)^n + \left(\frac{6}{2}\right)\left(\frac{4}{6}\right)^n - \left(\frac{6}{3}\right)\left(\frac{3}{6}\right)^n + \left(\frac{6}{4}\right)\left(\frac{2}{6}\right)^n - \left(\frac{6}{5}\right)\left(\frac{1}{6}\right)^n$
GEOMETRIC PROBABILITY

Let \( \Omega \) be a given region and \( A \) be a subset of \( \Omega \). We are interested in the probability that a randomly chosen point in \( \Omega \) falls in \( A \) or not. Here, randomly chosen means that a point may be any point of \( \Omega \) and that the probability of its falling in some subset \( A \) of \( \Omega \) is proportional to the measure of \( A \) (independent of the location or the shape of \( A \)).

Then the probability that a randomly chosen point in \( \Omega \) falls in \( A \), is defined \( \text{as: } P[A] = \frac{\text{Measure of } A}{\text{Measure of } \Omega} \).

Remark: Let \( X \) be a randomly chosen point in \( \Omega \). As \( X \) is chosen randomly, then the total probability unity is uniformly distributed over \( \Omega \), i.e.,

\[
\mathbb{P}(X = x) = \begin{cases} K \text{ (constant)} & \text{if } x \in \Omega, \\ 0 & \text{otherwise} \end{cases}
\]

Ex. 1. A point is picked at random from a unit square \( \Omega = \{(x, y) : 0 \leq x, y \leq 1\} \). Find the probability that a randomly chosen point in \( \Omega \) falls in

(a) \( A = \{(x, y) : \frac{1}{4} \leq x \leq \frac{3}{4}, \frac{1}{4} \leq y \leq \frac{3}{4}\} \)

(b) \( B = \{(x, y) : (x - \frac{1}{2})^2 + (y - \frac{1}{2})^2 = \frac{1}{4}\} \)

Solution: (a)

Required probability is

\[ P[A] = \frac{\text{Area of } A}{\text{Area of } \Omega} = \frac{\frac{1}{2} \times \frac{1}{2}}{1 \times 1} = \frac{1}{4}. \]

(b) Required probability is

\[ P[B] = \frac{\text{Area of } B}{\text{Area of } \Omega} = \frac{\pi \left(\frac{1}{2}\right)^2}{1 \times 1} = \frac{\pi}{4}. \]
Ex. 2: Two persons Amal and Bimal come to the club at random points of time between 6 PM and 7 PM, and stays for 10 minutes each. What is the chance that they will meet? [ISS'2012]

Solution:
Let \( X \): The time when Amal come to the club between 6 to 7 PM.
\( Y \): The time when Bimal come to the club between 6 to 7 PM.
Then, \( 0 \leq X, Y \leq 60 \), since 1hr. = 60 minutes.

The required probability is
\[
P \left[ |X - Y| \leq 10 \right] = \frac{\text{Area of the shaded region}}{\text{Area of the square}} = \frac{60 \times 60 - 50 \times 50}{60 \times 60} = 1 - \frac{25}{36} = \frac{11}{36},
\]

Remark: If they wait for 20 minutes and then they leave the place, then the required probability will be
\[
\frac{60 \times 60 - 40 \times 40}{60 \times 60} = \frac{5}{9},
\]
provided each of them comes at random to the spot during the specified time and train times of arriving are independent.

Ex. 3: \( n \) points are chosen at random and independently of one another inside a sphere of radius \( R \). Find the probability that the distance from the centre of the sphere to the nearest point is not less than \( R \) \(( < R \)).

Solution:
Let \( X_1, X_2, \ldots, X_n \) be the distances of the chosen \( n \) points from the centre of the sphere. Then, \( X_1, X_2, \ldots, X_n \) are independently distributed and have the same probability distribution.

Required probability = \( P \left[ X_1 > R, X_2 > R, \ldots, X_n > R \right] \)
\[
= P\left[ X_1 > R \right] \cdot P\left[ X_2 > R \right] \cdots P\left[ X_n > R \right] = \left[ P\left[ X_1 > R \right] \right]^n.
\]

Here, \( P\left[ X_1 > R \right] \) = The probability that a single chosen point lies on or outside the \( n \) smaller sphere.

Volume of the shaded region = \( \frac{4}{3} \pi (R^3 - r^3) = \left( 1 - \frac{r^3}{R^3} \right) \).

Hence, the required probability = \( \left( 1 - \frac{r^3}{R^3} \right)^n \).
Ex. 4. A bar of unit length is broken into three parts \(x, y, \frac{1}{2}\).

Find the probability that a triangle can be formed from the resulting parts. (OR)

Two points are chosen at random from a line segment. Show that the probability that the 2 points obtained this way from a triangle is \(\frac{1}{4}\).

**Solution:**

Three parts are: \(x, y, \frac{1}{2} = 1 - (x + y)\).

The conditions: \(x > 0, y > 0, x + y < 1\) are imposed on the quantities, \(x\) and \(y\), so, the sample space is the interior of a right

triangle with unit legs, so, Area of \(\Omega = \frac{1}{2}\).

The condition \(A\) requiring that a triangle could be formed from the segements \(x, y, \frac{1}{2} - (x + y)\) reduces to the following conditions:

1. The sum of any two sides is greater than the third side;
2. The difference between any two sides is smaller than third side.

This condition is associated with the triangle domain \(A\).

So, Area of \(A = \frac{1}{2} \times \frac{1}{4} = \frac{1}{8}\).

\[P[A] = \frac{\text{Area of } A}{\text{Area of } \Omega} = \frac{\frac{1}{8}}{\frac{1}{2}} = \frac{1}{4}\]

---

**REPEATED TRIALS**

- **Bernoulli Trial:** Repeated independent trials are said to be Bernoulli trials if each trial results in two outcomes, i.e., a success and a failure and the probability of success, \(p \ (0 < p < 1)\) remains same throughout the trials.

**Example:** Suppose a random experiment is repeated \(n\) times independently.

Then the occurrence of the event \(A\) may be termed as occurrence of a success and these repetitions constitute Bernoulli Trials.

Corresponding to Bernoulli trials, one may define \(n\) independent random variables \(X_1, X_2, \ldots, X_n\) \(\in \{0, 1\}\). \(P(X_i = 1) = p_1 = 1 - P(X_i = 0) = 1 - p_1\)

Here \(X_1, X_2, \ldots, X_n\) are called independent Bernoulli RV.

Clearly, PMF of \(X_i\) is given by

\[P(X_i = k) = p_i^{(k-1)}(1-p_i)^k, \quad 0 < k < 1, 0 < p_i < 1.\]
Ex.1. Two persons A and B toss a fair coin \((n+1)\) times and \(n\) times, respectively. Find the probability that 
(i) A will have as many heads as B 
(ii) A will have more heads than B

Solution: Let \(X\) and \(Y\) denote the number of heads obtained by A and B, respectively.

(i) \(P(X = Y) = P\left( \bigcup_{i=0}^{n} (X = i, Y = i) \right) = P\left[ \bigcup_{i=0}^{n} A_i \right]\), since \(A_i \cap A_j = \emptyset\) \(\forall i \neq j\)

\[
P(X = Y) = \sum_{i=0}^{n} P(A_i) = \sum_{i=0}^{n} P(X = i)P(Y = i)
\]

\[
= \sum_{i=1}^{n} \left( \frac{(n+1)}{2^{n+1}} \times \frac{n}{2^n} \right) = \frac{(2n+1)}{2^{2n+1}}.
\]

(ii) \(X = U+Y\)

\(U\): No of heads in 1st \(n\) tosses by A
\(V\): 1 on 0 according as the \((n+1)th\) toss by A results in heads or not.

\(X\) and \(Y\) are i.i.d. random variables.

\[
P(X > Y) = P(U+Y > Y)
\]

\[
= P(U+Y > Y | V=1)P(V=1) + P(U+Y > Y | V=0)P(V=0)
\]

\[
= \frac{1}{2} \left(P(U > Y-1) + P(U > Y)\right)
\]

\[
= \frac{1}{2} \left[P(U > Y) + P(U < Y)\right]
\]

\[
= \frac{1}{2} .
\]
Ex. 2. An urn contains $a$ white balls and $b$ black balls. 

(i) $n$ balls are chosen at random WR/WOR.

(ii) balls are drawn one by one WR/WOR till the $n$ white balls being produced.

Solution:

$X$: Number of white balls drawn

$Y$: Number of black balls preceding the $n^{th}$ white ball.

$Z$: Number of drawings required to produce $n$ white balls.

Clearly, $Z = n + Y$.

(i) $\text{WR:}$ $P(X = x) = \binom{n}{x} \left( \frac{a}{a+b} \right)^x \left( \frac{b}{a+b} \right)^{n-x}, \ x = 0, 1, 2, \ldots, n.$

(ii) $\text{WR:}$ $P(Y = y) = \binom{n-1+y}{y} \frac{a^n b y}{(a+b)^{n+y}}, \ y = 0, 1, 2, \ldots,$

$\text{WOR:}$ $P(Y = y) = \binom{n-1+y}{y} \left( \frac{a}{a+b} \right)^y \left( \frac{b}{a+b} \right)^{n+y}, \ y = 0, 1, 2, \ldots, b.$

$\text{WR:}$ $P(Z = z) = \binom{z-1}{n-1} \frac{a^n b^{z-n}}{(a+b)^z}, \ z = n, n+1, \ldots.$

$\text{WOR:}$ $P(Z = z) = \binom{z-1}{n-1} \frac{a^n (b)^{z-n}}{(a+b)^z}, \ z = n, n+1, \ldots, b+n.$
**RANDOM VARIABLES**

**Variables**

- **Random/Stochastic/Probabilistic**
- **Non-random/Degenerate random variables**

**Discrete**

**Continuous**

In any probability problem, we may associate with each outcome (elementary event) of the experiment of a finite real number. In many cases, the outcomes themselves are finite real numbers. This will be the case in tossing a die. In other cases, the numbers are artificially introduced, thus for example, in tossing a coin twice, the outcomes are not numbers but we may be interested in the number of heads obtained from the three tosses.

**Definitions of Random Variables:**

1. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space. Then a random variable is defined as a (Borel-measurable) function $X$ on the sample space $\Omega$ such that for every $\alpha \in \mathbb{R}$, the inverse image $X^{-1}((\alpha, \infty]) = \{\omega \in \Omega : X(\omega) \leq \alpha \}$ of the Borel set $(-\infty, \alpha]$ under $X$ is measurable with $\mathbb{P}( \alpha \in \mathcal{F})$.

2. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a sample space of a random experiment. A real valued function $X(\omega)$ defined on $\Omega$ is called a random variable if \[ \{ \omega : X(\omega) \leq \alpha \} \in \mathcal{F} \quad \forall \alpha \in \mathbb{R}. \]

3. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space of a random experiment. A finite single-valued function $X$ that maps $\Omega$ into $\mathbb{R}$ is called a random variable if the inverse image under $X$ of all Borel sets in $\mathbb{R}$ are events, i.e., if \[ X^{-1}(B) = \{ \omega : X(\omega) \in B \} \in \mathcal{F} \quad \forall B \in \mathcal{B}. \]

**Sample space $\Omega$**

- **Basic Prob. Measure $P$**

**Probability measure introduced by $X$ is called the prob. dist. of $X$.**
Although the induced probability measure $P_X(\cdot)$ for $\alpha \in \mathbb{R}^d$ instead of $B$ and also let $X$ is a random variable defined on a given probability space $(\Omega, \mathcal{A}, P)$ introduces the probability measure $P_X(\cdot)$. Note since $\{ \omega : -\infty < X(\omega) \leq \alpha \} = X^{-1}(\{-\infty, \alpha\}) \forall \alpha \in \mathbb{R}^d,$

$x \mapsto P_X([-\alpha, \alpha]) = P \{ \omega : -\alpha < X(\omega) \leq \alpha \} = F_X(\alpha), \ \alpha \in \mathbb{R}^d.$

Thus, for varying values of $\alpha \in \mathbb{R}^d$, the (point) function $F_X(\alpha)$ characterized the same as the (set) for $P_X([-\alpha, \alpha])$ does and accordingly is called the (cumulative) distribution function (d.f.) of the probability distribution of $X$.

Remark: (1) The notion of probability doesn't enter into the definition of a random variable.

(2) If $X$ is a random variable, the sets $\{X = \alpha\}$, $\{\alpha < X < \beta\}$, $\{X < \alpha\}$, $\{\alpha > X > \beta\}$, $\{\alpha \leq X \leq \beta\}$, etc. are all events. Moreover, one could have used any of these events to define a r.v.

Example of R.V.: 

(1) Let $E$: tossing of a fair coin. Then the sample space is $\Omega = \{H, T\}$. Let us define $X(H) = 1, \ X(T) = 0$. Then

$x^{-1}(-\alpha, \alpha) = \{ \omega : -\alpha < X(\omega) \leq \alpha \} = \begin{cases} \emptyset, & \text{if } \alpha < 0, \\ \{T\}, & \text{if } 0 \leq \alpha < 1 \\ \{H, T\}, & \text{if } 1 \leq \alpha. \end{cases}$

(2) Let $E$: tossing a coin twice. Then the sample space is $\Omega = \{HH, TH, HT, TT\}$. Define $X(\omega)$: the number of heads in $\omega$, $\omega \in \Omega$. Therefore, $X(HH) = 2, X(TH) = 1, = X(HT), X(TT) = 0$.

$x^{-1}(-\alpha, \alpha) = \{ \omega : -\alpha < X(\omega) \leq \alpha \} = \begin{cases} \emptyset, & \alpha < 0, \\ \{TT\}, & 0 \leq \alpha < 1 \\ \{TT, TH, HT\}, & 1 \leq \alpha \leq 2 \\ \Omega, & \alpha > 2. \end{cases}$

Hence, $X(\omega)$ is a random variable defined on $\Omega$. 
(3) Let $E$: tossing a coin three times.

$\Omega = \{HHH, HHT, HTH, HTT, THH, THT, TTH, TTT\}$

Define $X(\omega)$: the number of heads in $\omega$, for $\omega \in \Omega$.

Thus $X(\text{HHH}) = 3$, $X(\text{HHT}) = X(\text{THH}) = X(\text{HTH}) = 2$, $X(\text{TTT}) = 0$,

$X(\text{HTT}) = X(\text{THT}) = X(\text{TTT}) = 1$.

$X$ is a random variable with domain $\Omega$ and range $\{0, 1, 2, 3\}$

Here, $X^{-1}(a, \infty) = \begin{cases} 
\emptyset & \text{if } a < 0, \\
\{TTT\} & \text{if } 0 \leq a < 1, \\
\{HHT, THT, TTH\} & \text{if } 1 \leq a < 2, \\
\{HTT, HTH, THH\} & \text{if } 2 \leq a < 3, \\
\{HHH\} & \text{if } 3 \leq a < 4, \\
\Omega & \text{if } 4 \leq a.
\end{cases}$

Thus $X$ is a random variable here.

Here values of $X = \{3, 2, 2, 1, 2, 1, 1, 0\}$.

$X(\omega_1) = \begin{cases} 
0, \text{ if } i = 8, \\
1, \text{ if } i = 4, 6, 7, \\
2, \text{ if } i = 2, 3, 5, \\
3, \text{ if } i = 1.
\end{cases}$

For any particular event $\{X \leq 2.75\}$, the event space is $\{HHT, HTT, HTH, THT, TTH, THH, TTT\}$.

If $\{0.5 \leq X \leq 1.72\}$, then event space $= \{HHT, THT, TTH\}$.

(4) Let $E$: a coin is tossed until a head appears.

$X$: Number of tosses required.

Here $\Omega = \{H, T, TH, TTH, \ldots\}$ and $X$ assumes countably infinite number of values $1, 2, 3, \ldots$ corresponding $X(\omega_1) = 1, X(\omega_2) = 2$, etc.

Here $X^{-1}(a, \infty) = \begin{cases} 
\emptyset & \text{if } a < 1, \\
\{H\} & \text{if } 1 \leq a < 2, \\
\{TT\} & \text{if } 2 \leq a < 3, \\
\{TTH\} & \text{if } 3 \leq a < 4, \\
\ldots
\end{cases}$

Thus, $X$ is a random variable.
Problem 1: Let $X$ be a random variable, then
(a) Is $|X|$ also a random variable?
(b) Is $X^2$ also a random variable?

Solution:
Let $X$ be an r.v. defined on $(-\infty, \infty)$.

Then $\{\omega : X(\omega) \leq x\} \in \mathcal{F} \forall x \in \mathbb{R}$.

(a) Now, $|X(\omega)|$ is a real valued function defined on $(-\infty, \infty)$.
\[ \{\omega : |X(\omega)| \leq x\} = \{\omega : -x \leq X(\omega) \leq x\} \cap \{\omega : X(\omega) < -x\} \cap \{\omega : X(\omega) > x\} \cap \{\omega : X(\omega) = x\} \in \mathcal{F} \forall x \in \mathbb{R}. \]

Hence, $\{\omega : |X(\omega)| \leq x\} \in \mathcal{F} \forall x \in \mathbb{R}$.
So, $|X|$ is also a r.v. defined on $(-\infty, \infty)$.

(b) Clearly, $X^2(\omega)$ is a real valued function on $(-\infty, \infty)$.
\[ \{\omega : X^2(\omega) \leq x\} = \begin{cases} \emptyset & \text{if } x < 0 \\ \{\omega : -\sqrt{x} \leq X(\omega) \leq \sqrt{x}\} & \text{if } x > 0 \\ \emptyset & \text{if } x < 0 \\ \{\omega : X(\omega) \leq \sqrt{x}\} \cap \{\omega : X(\omega) < -\sqrt{x}\} & \text{if } x > 0 \end{cases} \in \mathcal{F} \forall x \in \mathbb{R}. \]

Hence, $X^2(\omega)$ is a random variable defined on $(-\infty, \infty)$.

Problem 2: If $X(\omega)$ is a random variable on $(-\infty, \infty)$, then show that $cX(\omega)$ is also a random variable on $(-\infty, \infty)$.

Proof:
Let $X$ be any arbitrary but fixed real number.

Then $(-\infty, x] \in \mathcal{F}$.

For $c > 0$,
\[ (cX)^{-1}(-\infty, x] = \{\omega : cX(\omega) \leq x\} = \{\omega : X(\omega) \leq \frac{x}{c}\} = X^{-1}(-\infty, \frac{x}{c}] \in \mathcal{F} \quad (\because X$ is an r.v.)

So, $cX(\omega)$ is also a random variable.
**DISTRIBUTION FUNCTION / CUMULATIVE DISTRIBUTION FUNCTION:**

**Definition:** Let \( X \) be a random variable defined on \((\Omega, \mathcal{A}, P)\). Define a point function \( F(\cdot) \) on \( \mathbb{R}^1 \) by

\[
F(x) = P\{w: X(w) \leq x\}, \quad \text{for all } x \in \mathbb{R}^1,
\]

is called the distribution function of R.V. \( X \).

**Properties:** (Alternative Definition*)

A real valued function \( F(x) \) defined on \( \mathbb{R} \) on \([-\infty, \infty] \) which satisfies the following properties:

1. \( \alpha_1 < \alpha_2 \implies F(\alpha_1) \leq F(\alpha_2) \) \( \forall \alpha_1, \alpha_2 \in \mathbb{R}^1 \).
   i.e., \( F(x) \) is monotonically non-decreasing.
2. \( F(-\infty) = \lim_{x \to -\infty} F(x) = 0 \).
3. \( F(\infty) = \lim_{x \to \infty} F(x) = 1 \).
4. \( F(x+0) = \lim_{h \to 0^+} F(x+h) = F(x) \) \( \forall x \in \mathbb{R}^1 \).
   i.e., \( F(x) \) is right continuous, is called a distribution function of \( X \).

**Proof of the properties of distribution function:**

1. \( \alpha_1 < \alpha_2 \)
   \( \implies \{ x \leq \alpha_1 \} \subseteq \{ x \leq \alpha_2 \} \)
   So, by the monotonicity theorem of probability,
   \[ P(X \leq \alpha_1) \leq P(X \leq \alpha_2) \]
   i.e., \( F(\alpha_1) \leq F(\alpha_2) \).
2. Let us take a sequence of events \( B_n = \{ X \leq -n \} \), \( n = 1, 2, \ldots \)
   \( \vdots B_n \) is a contracting sequence of events, i.e., monotonically decreasing. Hence, by \( \cap \) continuity theorem,
   \[ \lim_{n \to \infty} P(B_n) = P(\bigcap_{n \to \infty} B_n) \]
   \[ \lim_{n \to \infty} P(X \leq -n) = P(\bigcap_{n \to \infty} \{ X \leq -n \}) \]
   \[ \lim_{n \to \infty} P(X \leq -n) = P(\emptyset) = 0 \]
   \[ \lim_{n \to \infty} F(-n) = 0 \implies F(-\infty) = 0 \].
3. Let us take a sequence \( A_n = \{ X \leq n \} \)
   \( A_n \) is an expanding sequence of events, i.e., monotonically increasing. Hence, by \( \cup \) continuity theorem,
   \[ \lim_{n \to \infty} P(A_n) = P(\bigcup_{n \to \infty} A_n) \]
   \[ \lim_{n \to \infty} P(X \leq n) = P(\bigcup_{n \to \infty} (X \leq n)) \]
   \[ \lim_{n \to \infty} P(X \leq n) = P(\infty) = 1 \]
   \[ \lim_{n \to \infty} F(n) = F(\infty) = 1 \].
\[ \text{Let us take a sequence of events } \mathcal{C}_n = \{X \leq x + \frac{1}{n}\}, \quad n = 1, 2, \ldots \]

\[ \text{\(C_n\) is a contracting sequence of events, i.e., monotonically decreasing. Hence, by continuity theorem,} \]

\[ P(\lim_{n \to \infty} C_n) = \lim_{n \to \infty} P(C_n). \]

\[ \therefore P\left(\lim_{n \to \infty} \{X \leq x + \frac{1}{n}\}\right) = \lim_{n \to \infty} P(X \leq x + \frac{1}{n}) \]

\[ \text{i.e., } P(X \leq x) = \lim_{n \to \infty} P(X \leq x + \frac{1}{n}) \]

\[ \text{i.e., } F(x) = \lim_{n \to \infty} F(x + \frac{1}{n}) \]

\[ \text{Take, } \frac{1}{n} = h, \text{ as } n \to \infty, \ P_h \to 0. \]

\[ \lim_{n \to \infty} F(x + h) = F(x) \quad \text{on, } F(x + 0) = F(x). \]

\[ \text{Remark: } (1) \quad F(x) \text{ is not necessary continuous to the left.} \]

\[ \text{Justification: } \quad \text{Define, } D_n = \{\omega : X(\omega) \leq x - \frac{1}{n}\}, \quad n \in \mathbb{N} \]

\[ \text{Note that, } \lim_{n \to \infty} D_n = \lim_{n \to \infty} \{\omega : X(\omega) \leq x - \frac{1}{n}\} \]

\[ = \{\omega : X(\omega) < x\} \]

\[ \text{By continuity theorem of probability,} \]

\[ \lim_{n \to \infty} P[D_n] = P[\lim_{n \to \infty} D_n] \]

\[ \therefore \lim_{n \to \infty} P[\{\omega : X(\omega) \leq x - \frac{1}{n}\}] = P[\{\omega : X(\omega) < x\}] \]

\[ \lim_{n \to \infty} F(x - \frac{1}{n}) = P[\{\omega : X(\omega) \leq x\}] - P[\{\omega : X(\omega) = x\}] \]

\[ \lim_{h \to 0^+} F(x - h) = F(x) - P[X = x] \]

\[ F(x) - F(x - 0) = P[X = x] \geq 0 \]

\[ \text{Hence, } F(x - 0) \text{ is not necessary equal to } F(x), \text{ i.e.,} \]

\[ F(x) \text{ is not necessarily continuous to the left.} \]

\[ (2) \quad \text{Jump on saltus of a distribution function:} \]

\[ \text{If } P[X = a] = 0, \text{ then } F(a - 0) = F(a) \text{ and } F(x) \text{ is continuous} \]

\[ \text{at } x = a. \]

\[ \text{If } P[X = a] > 0, \text{ then the quantity } F(a) - F(a - 0) = P[X = a] \]

\[ \text{is called the jump on saltus of the d.f. } F(x) \text{ at } x = a. \]

\[ \text{Saltus/} \]

\[ \text{Jump} \]

\[ \begin{align*}
F(x) - F(x - 0) = & \quad P[X \leq x] - P[X < x] \\
F(x) = & \quad P[X \leq x] - P[X < x] \\
F(x - 0) = & \quad P[X = x] \\
\end{align*} \]

\[ \text{If } P[X = a] > 0, \text{ then } F(x) \text{ has discontinuity at } x = a \text{ with} \]

\[ \text{saltus } P[X = a]. \text{ So that the jump of a distribution function } F \]

\[ \text{at } x = a \text{ equals to the probability mass situated on concentrated} \]

\[ \text{at } x = a. \]
(3) A necessary and sufficient condition for the r.v. X on its d.f. F to be continuous at \( X = x \) is \( P[X = x] = 0 \).

**Proof:**
- Let \( P[X = x] = 0 \)
- Then \( F(x) = F(x-0) = 0 \)
- i.e. \( F(x) = F(x-0) \)........ (1)

Further, since, \( F \) is d.f., \( \therefore F(x) = F(x+0) \forall x \in \mathbb{R} \).

From (1) and (2), we have
- \( F(x) = F(x-0) = F(x+0) \)
- i.e. \( F \) is continuous at \( X = x \).

**Necessary:**
- \( F \) is continuous at \( X = x \).
- i.e. \( F(x) = F(x-0) = F(x+0) \)
- \( \Rightarrow F(x) - F(x-0) = 0 \)
- i.e. \( P[X = x] = 0 \).

\[ \therefore \text{The condition is necessary.} \]

---

**Ex. 1.** Let \( X \) be the r.v denoting "the number of heads in tossing a fair coin thrice". Find the cdf.

**Solution:**
- \( \Omega = \{ \text{HHH, HHT, HTH, THH, THT, HTT, TTT} \} \)

Note that \( X(\omega) = \begin{cases} 3 & \text{if } \omega = \text{HHH} \\ 2 & \text{if } \omega = \text{HHT, HTH, THH} \\ 1 & \text{if } \omega = \text{HTT, THT, HTT} \\ 0 & \text{if } \omega = \text{TTT} \end{cases} \)

Since the coin is fair, hence \( P[\omega] = \frac{1}{8} \), \( \forall \omega \in \Omega \).

The CDF of \( X \) is \( F(x) = P[X(\omega) \leq x] \)

**Diagram:**

**Note:** The set of values of \( X \) together with their corresponding probabilities is called the d.f. of \( X \).
Problem 1: Check whether the following function are distribution function on not:

\[ F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
\alpha & \text{if } 0 \leq x < \frac{1}{2} \\
1 & \text{if } x \geq \frac{1}{2}
\end{cases} \]

Solution:

(i) From the graph it is clear that the function is non-decreasing.
(ii) \( F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} 0 = 0 \).
(iii) \( F(\infty) = \lim_{x \to \infty} F(x) = \lim_{x \to \infty} 1 = 1 \).
(iv) \( \lim_{h \to 0} F(0+h) = \lim_{h \to 0} F(0) = F(0) \).
\( \lim_{h \to 0} F(\frac{1}{2}+h) = \lim_{h \to 0} (1) = 1 = F(\frac{1}{2}) \).
So, \( F(x) \) is right-continuous.
So, \( F(x) \) is a cdf here.

Problem 2: Is the following function cdfs on not?

\[ F(x) = \begin{cases} 
0 & \text{if } x \leq 0 \\
\frac{x^2}{2} & \text{if } 0 < x \leq 1 \\
\frac{1}{2} + \frac{(x-1)^3}{3} & \text{if } 1 < x \leq 2 \\
\frac{x^4}{7} + \frac{(x-2)^2}{3} & \text{if } 2 < x \leq 3 \\
1 & \text{if } x > 3
\end{cases} \]

Solution:

(i) It is non-decreasing.
(ii) \( F(-\infty) = 0 \), \( F(\infty) = 1 \).
(iii) \( F(\infty) = 1 \) if \( x > 3 \).
(iv) \( F(x+0) = \lim_{h \to 0^+} F(x+h) = \lim_{h \to 0^+} \frac{x^2}{2} = 0 = F(x) \).
\( \Rightarrow F(x) \) is continuous to the right at \( x = 0 \).
\( F(1+0) = \lim_{h \to 0^+} F(1+h) = \lim_{h \to 0^+} \left[ \frac{1}{2} + \frac{(1-h)^3}{3} \right] = \frac{1}{2} = F(1) \).
\( \Rightarrow F(x) \) is continuous to the right at \( x = 1 \).
\( F(2+0) = \lim_{h \to 0^+} F(2+h) = \lim_{h \to 0^+} \left[ \frac{2^4}{7} + \frac{(2-h)^2}{3} \right] = \frac{2^4}{7} = \frac{16}{7} \neq \frac{5}{3} = F(2) \).
\( \Rightarrow F(x) \) is not right continuous at \( x = 2 \).
\( \therefore F(x) \) can't be a cdf.
Problem: 3: The random variable $X$ assumes the value 'a' with probability unity. Sketch its $d.f.$.

Solution:

\[ P(X = a) = 1 \]
\[ P(X < a) = 0 \]
\[ P(X \leq a) = 1 \]
\[ P(X > a) = 1 - P(X \leq a) = 0 \]
\[ F(x) = \begin{cases} 0 & \text{if } x < a \\ 1 & \text{if } x > a \end{cases} \]

Problem: 4: The random variable $X$ assumes the value 1 with probability $p$ and the value 0 with probability $q$. Sketch the $d.f.$.

Solution:

\[ P(X = 0) = q, \quad P(X = 1) = p \]

So, $P(X \leq 1) = F(x)$

\[ F(x) = \begin{cases} 0 & \text{if } x < 0 \\ q & \text{if } 0 \leq x < 1 \\ 1 & \text{if } x \geq 1 \end{cases} \]

Since $F(0) = P(X \leq 0) = q$.

\[ F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = p + q = 1. \]

Problem: 5: A whole number is chosen at random between 1 and 10. Sketch the $d.f.$ of the related random variable.

Solution:

<table>
<thead>
<tr>
<th>$x$</th>
<th>$P(X = x)$</th>
<th>$F(x)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1/10$</td>
<td>$1/10$</td>
</tr>
<tr>
<td>2</td>
<td>$1/10$</td>
<td>$2/10$</td>
</tr>
<tr>
<td>3</td>
<td>$1/10$</td>
<td>$3/10$</td>
</tr>
<tr>
<td>4</td>
<td>$1/10$</td>
<td>$4/10$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>10</td>
<td>$1/10$</td>
<td>$1$</td>
</tr>
</tbody>
</table>
Ex. 5. Suppose \( G_1(x) = \begin{cases} 0 & \text{if } x < -1 \\ \frac{1}{2} & \text{if } -1 \leq x \leq 0 \\ a + b e^{-x^2/2} & \text{if } x > 0 \end{cases} \)

Determine the values of \( a \) and \( b \) so that \( G_1(x) \) is a distribution function.

Solution:

(i) \( G_1(-\infty) = 0 \)

(ii) \( G_1(\infty) = 1 \implies \lim_{x \to \infty} (a + b e^{-x^2/2}) = 1 \)

\[ a + b = 1 \implies a = 1. \]

(iii) \( G_1(x) \) is non-decreasing.

(iv) \( G_1(x) \) is right-continuous. So, we have

\[ G_1(0^+) = \frac{1}{2} = \lim_{h \to 0^+} G_1(0 + h) \]

\[ = \lim_{h \to 0^+} (a + b e^{-h^2/2}) \]

\[ = a + b \]

So, \( b = -\frac{1}{2} \).

Ex. 6. Verify whether the following function \( G_2(x) \) is a c.d.f. or not:

\[ G_2(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{1}{2} & \text{if } x = 0 \\ \frac{1}{2} + \frac{x}{2} & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases} \]

Solution:

(i) \( G_2(-\infty) = 0 \)

(ii) \( G_2(\infty) = 1 \)

(iii) From graph it is clear that \( G_2(x) \) is increasing.

(iv) \( G_2(x+0) = \lim_{h \to 0^+} F(0+h) \)

\[ = \lim_{h \to 0^+} \left( \frac{1}{2} + \frac{h}{2} \right) \]

\[ = \frac{1}{2} = G_2(0) \]

\[ = \frac{1}{2} = G_2(0) \]

\[ = \lim_{h \to 0^+} (1) \]

\[ = 1 = G_2(1) \]

So, \( G_2(x) \) is a c.d.f.
Ex. 7: Let $F_1$ and $F_2$ be two d.f.s. If $a$ and $b$ are non-negative integers whose sum is unity then show that $aF_1 + bF_2$ is also d.f.s.

**Solution:**

(i) $F = aF_1 + bF_2$.

Let $\alpha_1 < \alpha_2$.

Then since $F_1$ and $F_2$ are d.f.s, so, we have

$F_1(\alpha_1) \leq F_1(\alpha_2)$ and $F_2(\alpha_1) \leq F_2(\alpha_2)$.

Since $a$ and $b$ are non-negative integers, so

$aF_1(\alpha_1) + bF_2(\alpha_1) \leq aF_1(\alpha_2) + bF_2(\alpha_2)$

$\implies (aF_1 + bF_2)(\alpha_1) \leq (aF_1 + bF_2)(\alpha_2)$

$\implies F(\alpha_1) \leq F(\alpha_2)$. So, $F$ is non-decreasing.

(ii) $F_1(-\infty) = 0$, $F_2(-\infty) = 0$.

$F(-\infty) = (aF_1 + bF_2)(-\infty)$

$= aF_1(-\infty) + bF_2(-\infty)$

$= 0$.

(iii) $F_1(\infty) = 1$, $F_2(\infty) = 1$.

$F(\infty) = (aF_1 + bF_2)(\infty)$

$= aF_1(\infty) + bF_2(\infty)$

$= a + b = 1$.

(iv) Now, $F_1(\alpha + 0) = \lim_{h \to 0} F_1(\alpha + h) = F_1(\alpha)$ \forall $\alpha$.

And, $F_2(\alpha + 0) = \lim_{h \to 0} F_2(\alpha + h) = F_2(\alpha)$ \forall $\alpha$.

Now, $F(\alpha + 0) = \lim_{h \to 0} (aF_1 + bF_2)(\alpha + h)$

$= \lim_{h \to 0} aF_1(\alpha + h) + bF_2(\alpha + h)$

$= aF_1(\alpha) + bF_2(\alpha)$

$= F(\alpha)$ \forall $\alpha$.

Hence, $F(\alpha)$ is right continuous.

So, $F$ is a d.f.
Ex. 8. \( F(x) \) is a d.f. Then show that \( G_n(x) \) is also a d.f., where

\[
G_n(x) = \left[ 1 - (1 - F(x))^n \right], \quad n \in \mathbb{N}
\]

**Solution:**

(i) \( \alpha < \beta \)

\[
F(\alpha) \leq F(\beta)
\]

\[
\Rightarrow 1 - F(\alpha) > 1 - F(\beta)
\]

\[
\Rightarrow (1 - F(\alpha))^n > (1 - F(\beta))^n
\]

\[
\Rightarrow 1 - (1 - F(\alpha))^n \leq 1 - (1 - F(\beta))^n
\]

\[
G_n(\alpha) \leq G_n(\beta)
\]

(ii) \( G_n(-\infty) = \lim_{x \to -\infty} G_n(x) \)

\[
= \lim_{x \to -\infty} \left[ 1 - (1 - F(x))^n \right]
\]

\[
= 1 - \lim_{x \to -\infty} (1 - F(x))^n
\]

\[
= 1 - \left( 1 - \lim_{x \to -\infty} F(x) \right)^n
\]

\[
= 1 - (1 - \lim_{x \to -\infty} F(x))^n
\]

\[
= 1 - 1 = 0
\]

(iii) \( G_n(\infty) = \lim_{x \to \infty} G_n(x) \)

\[
= \lim_{x \to \infty} \left[ 1 - (1 - F(x))^n \right]
\]

\[
= 1 - \left( 1 - \lim_{x \to \infty} F(x) \right)^n
\]

\[
= 1 - (1 - F(\infty))^n
\]

\[
= 1
\]

(iv) \( \lim_{h \to 0} G_n(x+h) \)

\[
= \lim_{h \to 0} \left[ 1 - (1 - F(x+h))^n \right]
\]

\[
= 1 - \left( 1 - \lim_{h \to 0} F(x+h) \right)^n
\]

\[
= 1 - (1 - F(x))^n
\]

\[
= G_n(x)
\]

\( \therefore G_n(x) \) is also a d.f.
Ex. 9. Show that every d.f. F has the following properties:

(a) \( \lim_{\alpha \to \infty} \alpha \int_{-\infty}^{\alpha} \frac{1}{2} dF(\alpha) = 0 \)

(b) \( \lim_{\alpha \to 0} \alpha \int_{-\alpha}^{0} \frac{1}{2} dF(\alpha) = 0 \)

(c) \( \lim_{\alpha \to -\infty} \alpha \int_{-\infty}^{-\alpha} \frac{1}{2} dF(\alpha) = 0 \)

(d) \( \lim_{\alpha \to -\infty} \alpha \int_{0}^{\alpha} \frac{1}{2} dF(\alpha) = 0 \)

Solution:

(b) \( \) Let \( \alpha \) be a positive proper fraction.

\[
0 \leq \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) = \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) + \alpha \int_{0}^{\frac{1}{\alpha}} \frac{1}{2} dF(\alpha)
\]

For, \( \alpha \to 0^+ \), first part \( \to 0 \).

\[
1 < \frac{1}{\alpha} \to 0 < \frac{1}{\alpha} \to 1.
\]

\[
0 \leq \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) \leq \alpha \int_{-\infty}^{\frac{1}{\alpha}} \frac{1}{2} dF(\alpha)
\]

\[
\leq \alpha \int_{-\infty}^{\frac{1}{\alpha}} \frac{1}{2} dF(\alpha)
\]

\[
\leq \sqrt{\alpha} \left[ \cdot 1 - F(\frac{1}{\alpha}) \right] < 1
\]

\[
\lim_{\alpha \to 0^+} \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) = 0.
\]

(c) \( \alpha < 0 \) and \( 0 \leq \alpha \leq 0 \).

\[
\frac{\alpha}{2} \leq 1.
\]

\[
0 \leq \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) \leq \int_{-\infty}^{0} dF(\alpha) = F(\alpha)
\]

\[
\alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) \leq F(\alpha)
\]

\[
\Rightarrow 0 \leq \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) \leq \lim_{\alpha \to -\infty} F(\alpha)
\]

\[
\Rightarrow 0 \leq \lim_{\alpha \to -\infty} \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) = 0.
\]

\[
\lim_{\alpha \to -\infty} \alpha \int_{-\infty}^{0} \frac{1}{2} dF(\alpha) = 0.
\]
Ex. 10. If $X_1$ and $X_2$ are independently and identically distributed N(0, 2), prove that:

(i) $P\left\{ |X_1 - X_2| > t \right\} \leq 2P \left\{ |X_1| > \frac{t}{2} \right\}$.

(ii) If $a > 0$ such that $P\left\{ X_1 \geq a \right\} = 1 - p$, then $P\left\{ |X_1 - X_2| > t \right\} \geq pP\{ |X_1| > a + t \}$, for $t > 0$.

Solution:

(i) $\{ |X_1| > \frac{t}{2} \} \cup \{ |X_2| > \frac{t}{2} \} \supseteq \{ |X_1 - X_2| > t \}$

$\therefore P\left\{ |X_1 - X_2| > t \right\} \leq P\{ |X_1| > \frac{t}{2} \} + P\{ |X_2| > \frac{t}{2} \}$

$= 2P \left\{ |X_1| > \frac{t}{2} \right\}$.

(ii) $\{X_1 > a+t, X_2 \leq a\} \cup \{X_1 \leq -a-t, X_2 \geq -a\} \subset \{ |X_1 - X_2| > t \}$

$\therefore P\left\{ |X_1 - X_2| > t \right\} \geq P\{ X_1 > a+t, X_2 \leq a \} + P\{ X_1 \leq -a-t, X_2 \geq -a \}$

$\geq P\{ X_1 > a+t \} + P\{ X_1 \leq -a-t \}$

Alternative proof of (i):

Note that, $|X_1| \leq \frac{t}{2}, |X_2| \leq \frac{t}{2} \Rightarrow |X_1 - X_2| \leq t$.

From the monotonicity theorem of probability,

$P\{ |X_1| \leq \frac{t}{2}, |X_2| \leq \frac{t}{2} \} \leq P\{ |X_1 - X_2| \leq t \}$

But $X$ and $Y$ are i.i.d. N(0, 2), so $P\{ |X_1| \leq \frac{t}{2} \} = P\{ |X_2| \leq \frac{t}{2} \}$

So, $P\{ |X_1 - X_2| > t \} \leq 2P\{ |X_1| > \frac{t}{2} \}$. 
A. Discrete Random Variable:

Definition: A random variable $X$ takes only a countable (finite or infinite) number of isolated values $x_1, x_2, \ldots, x_n, \ldots$ with $\Pr[X = x_i] > 0$ for all $i$, is called a discrete random variable.

The points $x_1, x_2, \ldots$ that have positive probabilities of occurrence are called the jump points of the r.v. $X$.

Probability Mass Function: Let $X$ be a discrete r.v. with mass points $x_1, x_2, \ldots$. Then $\Omega = \bigcup_{i=1}^{\infty} \{\omega: X(\omega) = x_i\}$ and

$$\Pr[\Omega] = \sum_{i=1}^{\infty} \Pr[X = x_i] \quad \text{[By countable additivity of $\Pr[\cdot]$]}
$$

Definition: Let $X$ be a discrete r.v. with mass points $\{x_1, x_2, \ldots\}$, then the function

$$f(x) = \begin{cases} \Pr[X = x_i] & \text{if } x = x_i, \ i = 1, 2, \ldots \\ 0 & \text{if } x \neq x_i \end{cases}
$$

is called the PMF of the r.v. $X$.

Theorem: A function $f(x)$ is said to be a PMF of some discrete r.v. $X$ if

(i) $f(x) > 0$ for $x \in \mathbb{R}$

(ii) $\sum_x f(x) = 1$.

Alternative Definition: The probability mass function $f(x)$ of a r.v. $X$ whose set of possible values are $\{x_1, x_2, \ldots\}$ is a function from $\mathbb{R}$ to $\mathbb{R}$ that satisfies the following properties:

(i) $f(x) = 0$ if $x \neq x_i$

(ii) $f(x) = \Pr[X = x_i]$ if $x = x_i, \ i = 1, 2, \ldots$

(iii) $\sum_x f(x) = 1$.

Given the pmf of a discrete distribution, we can get the distribution function by successive addition, i.e.,

$$F(x) = f(x_1) + f(x_2) + f(x_3) + \ldots + f(x_n),$$

where $x_1 < x_2 < x_3 < \ldots < x_n < \ldots \ x_i$.

On the other hand, given the df, we can get the pmf by successive subtraction, i.e.,

$$f(x) = F(x) - F(x-1) = \Pr[X \leq x] - \Pr[X < x].$$

is the probability at the point $x_i$. 
Ex. 1. For what values of $\theta$ and $c$ is the function $f$ given by

$$f(x) = \begin{cases} \frac{c \cdot \theta^x}{x}, & x=1,2,3, \ldots, \\ 0, & \text{otherwise} \end{cases}$$

a PMF?

Solution:

(i) As $f(x) > 0$ for $x=1,2,3, \ldots$, hence $c > 0$, $\theta > 0$.

(ii) $1 = \sum_{x} f(x) = \sum_{x=1}^{\infty} \frac{c \cdot \theta^x}{x}$

$$= c \sum_{x=1}^{\infty} \frac{\theta^x}{x}$$

$$= c \{ - \log_e (1-\theta) \}^\theta, \text{ if } 0 < \theta < 1.$$

$$\therefore c = - \frac{1}{\log_e (1-\theta)}$$

and $0 < \theta < 1$.

Ex. 2. Let $f(x) = \begin{cases} p \cdot q^x, & x=0,1,2,3, \ldots; p+q=1, \text{ or } 0 < q < 1. \\ 0, & \text{otherwise} \end{cases}$

Does $f(x)$ define a PMF of some RV $X$? What is the DF of $x$?

Find $P[n \leq X \leq m]$, $n, m \in \mathbb{N}$.

Solution:

(i) $0 < q < 1$ \[ \rightarrow 1-q > 0 \]

\[ \rightarrow (1-q)^x > 0 \quad \therefore x=0,1,2,3, \ldots. \]

\[ \rightarrow p \cdot (1-q)^x > 0 \quad \therefore 0 < q < 1. \]

\[ \therefore f(x) > 0. \]

(ii) $\sum_{x=0}^{\infty} f(x) = \sum_{x=0}^{\infty} p \cdot q^x = \sum_{x=0}^{\infty} q^x = \frac{p}{1-q} = \frac{1-q}{p} = 1.$

\[ \therefore f(x) \text{ defines a PMF of some RV } X. \]

(iii) $F(x) = P[X \leq x] = p + q^1 + q^2 + \ldots + q^x$

\[ = p \left[ 1 + q + q^2 + \ldots + q^x \right] \]

\[ = p \cdot \frac{1-q^{x+1}}{1-q} \quad \therefore x=0,1,2, \ldots. \]

\[ = 1-q^{x+1}; \]

(iv) $P[n \leq X \leq m] = P[X \leq m] - P[X < n]$

\[ = P[X \leq m] - P[X \leq n-1] \]

\[ = F(m) - F(n-1) = \{ 1 - q^{m+1} \} - \{ 1 - q^n \} \]

\[ = (q^n - q^{m+1}). \]
Ex. 3. Find the PMF of the RY X whose DF is

\[ F(x) = \begin{cases} 
0 & , \ x < 0 \\
\frac{i(i+1)}{n(n+1)} & , \ i \leq x < i+1, \ i = 0, 1, \ldots, (n-1), \\
1 & , \ x \geq n. 
\end{cases} \]

Solution:
Note that \( i = 1, 2, 3, \ldots, n. \)

\[ P[X=i] = P[X \leq i] - P[X < i] \]
\[ = F(i) - F(i-1) \]
\[ = \frac{i(i+1)}{n(n+1)} - \frac{(i-1)i}{n(n+1)} \]
\[ = \frac{2i}{n(n+1)}. \]

The PMF of \( X \) is \( f(x) = \begin{cases} \frac{2x}{n(n+1)} & , \ x = 1(1)n, \\
0 & , \ \text{otherwise}. \end{cases} \]

**Distribution Function of Discrete Random Variables**

Let \( X \) be a discrete R.V. with mass points \( \alpha_1 < \alpha_2 < \ldots \ldots \) Then the D.F. is

\[ F(x) = P[X \leq \alpha] = \begin{cases} 
0 & , \ x < \alpha_1 \\
\sum_{i=1}^{\alpha_2} P[X = \alpha_i] & , \ \alpha_1 \leq x < \alpha_2 \\
\sum_{i=1}^{\alpha_3} P[X = \alpha_i] & , \ \alpha_2 \leq x < \alpha_3 \\
\vdots \\
\sum_{i=1}^{\alpha_k} P[X = \alpha_i] & , \ \alpha_{k-1} \leq x < \alpha_k, \ k = 1, 2, 3, \ldots. 
\end{cases} \]

Hence, the D.F. \( F(x) \) of a discrete R.V. has discontinuity points as the mass points of the R.V. The number of discontinuity points is the same as the no. of mass points.
Ex. 6. Can a function of the form \( f(x) = \begin{cases} \left( \frac{2}{3} \right)^x & \alpha = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases} \) be a probability mass function (PMF)?

Solution:

Note that, since \( f(x) > 0 \) if \( \alpha > 0 \).

And, to be a p.d.f., the below condition also needs to be satisfied:

\[
\sum_{i=1}^{\infty} c \left( \frac{2}{3} \right)^i = 1
\]

\[\Rightarrow c \cdot \frac{\frac{2}{3}}{1 - \frac{2}{3}} = 1\]

\[\Rightarrow c = \frac{1}{2}\]

Thus, only for \( c = \frac{1}{2} \), \( f(x) \) can be a PMF.

Ex. 7. Let \( X \) be the number of births in a hospital until the first girl is born. Assume that the probability is \( \frac{1}{2} \) that a baby born is a girl. Determine the PMF and DF of \( X \).

Solution:

\( X \) is an n.v. that can assume any positive integer \( i \), \( f(i) = P(X = i) \), and \( X \geq 1 \) occurs if the first \( i-1 \) births are all boys and the \( i \text{th} \) birth is a girl.

Thus \( f(i) = \left( \frac{1}{2} \right)^{i-1} \left( \frac{1}{2} \right) = \left( \frac{1}{2} \right)^i \) for \( i = 1, 2, 3, \ldots \)

and \( f(x) = 0 \) if \( x \neq 1, 2, 3, \ldots \)

\( F(t) = \begin{cases} \frac{1}{2} & \text{if } 1 \leq t < 2 \\ \frac{1}{2} + \frac{1}{4} & \text{if } 2 \leq t < 3 \\ \frac{1}{2} + \frac{1}{4} + \frac{1}{8} & \text{if } 3 \leq t < 4 \\ \vdots \\ \frac{1}{2} + \frac{1}{2^2} + \cdots + \frac{1}{2^{n-1}} & \text{if } (n-1) \leq t < n \end{cases} \)

So, \( F(t) = \begin{cases} \frac{n-1}{2} & \text{if } t < 1 \\ \sum_{i=1}^{n-1} \left( \frac{1}{2} \right)^i & \text{if } n-1 \leq t < n, n = 2, 3, 4, \ldots \\ 0 & \text{if } t < 1 \\ 1 - \left( \frac{1}{2} \right)^{n-1} & \text{if } n-1 \leq t < n \forall n = 2, 3, 4, \ldots \end{cases} \)
Continuous Random Variable:

Definition: A random variable $X$ is said to be a continuous RV if it takes any value within its range of variation.

For a continuous RV $X$, $P[X = x] = 0 \ \forall x$.

By construction or axiomatic definition,

$$F(x) - F(x-0) = P[X = x] = 0 \ \forall x.$$  

$\Rightarrow F(x)$ is continuous everywhere.

If $F(x)$ is continuous everywhere, then the associated RV $X$ is known as a Continuous Random Variable.

Absolutely continuous Random Variable: An RV $X$ with D.F. $F(x)$ is said to be an absolutely continuous RV, if $\exists$ a non-negative function $f(x)$ such that

$$F(x) = \int_{-\infty}^{x} f(t) dt, \ \forall x \in \mathbb{R}.$$  

where $F(x) = P[X \leq x]$ is the distribution function of the RV $X$.

It may be noted that

(i) $F(-\infty) = \lim_{x \to -\infty} F(x) = \lim_{x \to -\infty} \int_{-\infty}^{x} f(x) dx = 0.$

(ii) $F(\infty) = \lim_{x \to \infty} F(x) = \lim_{x \to \infty} \int_{-\infty}^{x} f(x) dx = 1.$

(iii) $P[a < x \leq b] = F(b) - F(a)$

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

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

$$= \int_{a}^{b} f(x) dx = P[a < x \leq b] = P[a \leq x < b] = P[a \leq x \leq b].$$

And the function $f(x)$ is called the probability density function (pdf).

Theorem: A function $f(x)$ is said to be a PDF of some absolutely continuous RV $X$ if it satisfies

(i) $f(x) \geq 0 \ \forall x$

(ii) $\int_{-\infty}^{\infty} f(x) dx = 1.$
Result: If \( F(x) \) is absolutely continuous and \( f(x) \) is continuous at 
\( x = a \), then 
\[
F'(x) = \frac{dF(x)}{dx} = f(x).
\]

Proof:

Necoton-Liebnitz Formula:

\[
I(a) = \int_{b(a)}^{a} f(x, \theta) \, dx \quad \text{then} \quad I'(a) \quad \text{is defined as} \quad \frac{dI(a)}{d\theta} = \int_{b(a)}^{a} \frac{df}{d\theta} \, dx + \int_{a}^{b(a)} \frac{db}{d\theta} f(b(a), \theta) - \frac{d\theta}{d\theta} f(a(a), \theta)
\]

Here, 
\[
F'(x) = \frac{dF(x)}{dx} = \frac{d}{dx} \int_{-\infty}^{x} f(x) \, dx = \int_{-\infty}^{x} \frac{df(x)}{dx} \, dx + \int_{-\infty}^{x} f(x) \, dx = \int_{-\infty}^{x} f(x) \, dx + p(x) = f(x).
\]

Probability Density Function (PDF):

For an absolutely continuous RV 
\( X \) with D.F. \( F(x) \), note that

\[
\frac{d}{dx} F(x) = f(x),
\]

\[
\Rightarrow f(x) = \lim_{h \to 0^+} \frac{F(x+h) - F(x)}{h} = \lim_{h \to 0^+} \frac{P[x \leq x \leq x+h]}{h}, \quad \text{since} \quad P[x = x] = 0.
\]

For small \( h > 0 \), \( f(x) \approx \frac{P[x \leq x \leq x+h]}{h} \), which is the ratio of the probability contained in \( (x, x+h] \) for the distribution and the length of the interval, i.e., \( f(x) \approx \frac{P[x \leq x \leq x+h]}{h} \) is the probability contained for the distribution per unit length in the interval \( (x, x+h] \), where \( h > 0 \) is small.

That is, the quantity \( f(x) \) is known as the probability density at the point \( x \) and the function \( f(x) \) is called pdf of RV \( X \).

Definition: If \( X \) is an absolutely continuous RV \( X \) with D.F. \( F(x) \), 

\[
\exists \text{a non-negative function } f(x) \quad \text{such that}
\]

\[
F(x) = \int_{x}^{\infty} f(t) \, dt \quad \text{for} \quad x \in \mathbb{R}
\]

is called the pdf of \( X \). It satisfies the properties:

(i) \( f(x) \geq 0 \quad \forall \quad x \in \mathbb{R} \)

(ii) \( \int_{-\infty}^{\infty} f(t) \, dt = 1 \).
Question: Why is called a PDF? Justify your answer.

Solution:

\[ \int_{x}^{x+dx} f(x) \, dx = \int_{x}^{x+dx} F'(x) \, dx = \left[ F(x) \right]_{x}^{x+dx} \]

\[ = F(x+dx) - F(x) \]

\[ = P \left[ x < X \leq x + dx \right] \]

\[ = \lim_{\Delta x \to 0} F \left[ x < X \leq x + \Delta x \right] \]

\[ = \lim_{\Delta x \to 0} \left\{ \frac{F(x+4x) - F(x)}{\Delta x} \right\} \Delta x \]

\[ = \lim_{\Delta x \to 0} \left\{ \frac{F(x+4x) - F(x)}{\Delta x} \right\} \lim_{\Delta x \to 0} \Delta x \]

\[ = F'(x) \, dx \]

\[ \Rightarrow \frac{dF}{dx} = f(x) \, dx, \text{ where } f(x) \text{ is the PDF of } X. \]

In probability theory, probability has been regarded as "mass" and "dx" is the one-dimensional analogue of "volume," so that \( dP \) plays the role of "mass(m)" and "dx" plays the role of "volume(V)" in the relation \( M = P \cdot V \) of elementary physics.

Therefore, the function \( f(x) \) is playing the role of "density," and that is why \( f(x) \) is called the probability density function within an infinitesimal interval about \( x = x \).

Note: For non- absolutely continuous RV, the PDF doesn't exist.
Ex. 1. Verify that the function \( f(x) \) can be looked upon as the PDF of a continuous random variable.

\[
f(x) = \begin{cases} 
\frac{x}{2} & , \ 0 < x \leq 1 \\
\frac{1}{2} & , \ 1 < x \leq 2 \\
\frac{3-x}{2} & , \ 2 < x \leq 3 \\
0 & , \ 3 < x \leq 4 
\end{cases}
\]

Obtain the distribution function.

Solution: Clearly, \( f(x) \geq 0 \ \forall \ x \in \mathbb{R} \).

\[
\int_{-\infty}^{\infty} f(x) \, dx = \int_{-\infty}^{0} 0 \, dx + \int_{0}^{1} \frac{x}{2} \, dx + \int_{1}^{2} \frac{1}{2} \, dx + \int_{2}^{3} \frac{3-x}{2} \, dx + \int_{3}^{\infty} 0 \, dx = 1.
\]

Hence, \( f(x) \) is a PDF.

The D.F. is

\[
F(x) = \int_{-\infty}^{x} f(t) \, dt
\]

\[
= \begin{cases} 
0 & , \ x < 0 \\
\frac{x^2}{4} & , \ 0 < x \leq 1 \\
\frac{x^2}{4} + \frac{x}{2} & , \ 1 < x \leq 2 \\
\frac{x^2}{4} + \frac{x}{2} + \left( \frac{3-x}{2} \right) & , \ 2 < x \leq 3 \\
\frac{6x-x^2}{4} & , \ 3 < x \leq 4 \\
1 & , \ x > 4
\end{cases}
\]

Ex. 2. Let \( f(x) = \begin{cases} 
K & , \ 0 < x < \frac{1}{2} \\
0 & , \ \text{otherwise}
\end{cases} \) be a PDF of \( x \). Find the constant \( k \).

Solution: \( f(x) > 0 \ \Rightarrow \ K > 0 \).

\[
\int_{-\infty}^{\infty} f(x) \, dx = \int_{0}^{\frac{1}{2}} K \, dx = K \cdot \frac{1}{2} = 1 \ \Rightarrow \ K = 2.
\]
Ex. 10. If \( X \) has an absolutely continuous distribution with PDF \( f(x) \) shown below, then find \( \mu \) and \( \sigma \).

(i) \( f(x) = \frac{\theta}{2} e^{-\theta |x - a|} \), where \( \theta > 0 \)

(ii) \( f(x) = \frac{a}{\pi \left[ a^2 + (x - \theta)^2 \right]} \), where \( a > 0 \).

Solution:

\[ f(x) = \begin{cases} \frac{\theta}{2} e^{-\theta (x - a)} & \text{if } x > a \\
\frac{\theta}{2} e^{-\theta (a - x)} & \text{if } x \leq a \end{cases} \]

Now, \( F(x) = \begin{cases} 
\frac{\theta}{2} e^{-\theta (x - a)} & \text{if } x \leq a \\
\frac{1}{\pi} \arctan \left( \frac{x - \theta}{a} \right) & \text{if } x > a \end{cases} \)

\( F(x) = \int_{-\infty}^{x} f(t) \, dt = \frac{a}{\pi} \int_{-\infty}^{x} \frac{1}{\left[ a^2 + (t - \theta)^2 \right]} \, dt \)

\[ = \frac{a}{\pi} \left[ \frac{1}{a} \arctan \left( \frac{t - \theta}{a} \right) \right]_{-\infty}^{x} \]

\[ = \frac{1}{\pi} \left[ \arctan \left( \frac{x - \theta}{a} \right) + \frac{\pi}{2} \right] \text{ if } a > 0. \]

So, \( F(x) = \begin{cases} 
\frac{1}{\pi} \arctan \left( \frac{x - \theta}{a} \right) + \frac{1}{2} & \text{if } a > 0 \\
0 & \text{if } x \leq 0 \end{cases} \)
Ex.11. (a) Sketch the graph of the function
\[ f(x) = \begin{cases} \frac{1}{2} - \frac{1}{4}|x-3| & \text{if } 1 \leq x \leq 5 \\ 0 & \text{Otherwise} \end{cases} \]
and show that it is the PDF of an RV $X$.
(b) Find $F_x(t)$, the d.f. of $X$, and show that it is continuous.
(c) Sketch the graph of $F_x(t)$.

Solution: (a) $f(x) = \begin{cases} 0 & \text{if } x < 1 \\ \frac{1}{2} + \frac{1}{4}(x-3) & \text{if } 1 \leq x < 3 \\ \frac{1}{2} - \frac{1}{4}(x-3) & \text{if } 3 \leq x < 5 \\ 0 & \text{if } x \geq 5 \end{cases}$

Graph of PDF:

Since $f(x) > 0$ and the area under $f$ from 1 to 5, being the area of the triangle $ABC$ is $\frac{1}{2} \times 4 \times 1 = 1$, so $f$ is a density function of some random variable $X$.

(b) $F_x(t) = \begin{cases} 0 & \text{if } t < 1 \\ \int_{1}^{t} \left[ \frac{1}{2} + \frac{1}{4}(x-3) \right] dx = \frac{1}{8}t^2 - \frac{1}{4}t + \frac{1}{6} & \text{if } 1 \leq t < 3 \\ \int_{1}^{3} \left[ \frac{1}{2} + \frac{1}{4}(x-3) \right] dx + \int_{3}^{t} \left[ \frac{1}{2} - \frac{1}{4}(x-3) \right] dx = -\frac{1}{8}t^2 + \frac{5}{4}t - \frac{17}{8} & \text{if } 3 \leq t < 5 \\ \int_{1}^{3} \left[ \frac{1}{2} + \frac{1}{4}(x-3) \right] dx + \int_{3}^{5} \left[ \frac{1}{2} - \frac{1}{4}(x-3) \right] dx = 1 & \text{if } t \geq 5 \end{cases}$

$F_x(t)$ is continuous because $\lim_{t \to 1^-} F_x(t) = 0 = F_x(1)$, $\lim_{t \to 3^-} F_x(t) = \frac{1}{8}(3)^2 - \frac{1}{4}(3) + \frac{1}{6} = F_x(3)$. 

(c) Graph of $F_x(t)$.
EX. 16. A fair die is thrown repeatedly till every face has appeared at least once. Let $X$ denote the number of throws made. Find the distribution of $X$.

Solution:

$$P(X = x) = P(X = x) - P(X = x - 1)$$

$$= \left( \text{The probability that in } x \text{ throws all varieties will appear} \right) - \left( \text{The probability that in } (x-1) \text{ throws all varieties will appear} \right)$$

$$= \left\{ 1 - \left( \frac{1}{6} \right) \left( \frac{5}{6} \right)^{x} + \left( \frac{1}{6} \right) \left( \frac{5}{6} \right)^{x-1} \right\}$$

$$- \left\{ 1 - \left( \frac{1}{6} \right) \left( \frac{5}{6} \right)^{x-1} + \left( \frac{1}{6} \right) \left( \frac{5}{6} \right)^{x-2} \right\}$$

EX. 17. An urn contains $N$ cards labelled from 1 to $N$. If $n$ drawing are made at random without replacements from the urn, let $X$ denote the least number drawn. Find the distribution of the random variable $X$.

If $Y$ denote the highest number drawn, obtain the distr of $Y$.

Solution: $X$ denotes the least number drawn. The mass points of $X$ are $1, 2, \ldots, N-n+1$. Now $n$ cards can be drawn out of $N$ cards in $n$ ways and in order that the least number of cards drawn is $x$, we have to select the $n-1$ cards other than drawn is $x$, we have $\binom{x-1}{n-1}$ ways to select the $n-1$ cards from the set $x+1, x+2, \ldots, n$, i.e.,

$$\text{from } N-x \text{ cards and this can be done in } \binom{N-x}{n-1} \text{ ways.}$$

Hence, $P(X = x) = \frac{\binom{N-x}{n-1}}{\binom{N}{n}}$, $x = 1, 2, \ldots, N-n+1$.

$Y$ denotes the highest number drawn. The mass points of $Y$ are $n, n+1, \ldots, N$.

Total number of ways in which $n$ cards can be drawn from the set is $\binom{N}{n}$. In order that the highest number drawn is $y$, we have to select remaining $(n-1)$ cards from the card number $y$ from the set $1, 2, 3, \ldots, y-1$, i.e., from $(y-1)$ cards and this can be done in $\binom{y-1}{n-1}$ ways.

Hence, $P(Y = y) = \frac{\binom{y-1}{n-1}}{\binom{N}{n}}$, $y = n, n+1, \ldots, N$. 

Scanned by CamScanner
Ex. 18. A man counts to open his door and has $n$ keys, only one of which fits the door. For some reason which can only be surmised, he tries the keys independently and at random. Find a probability distribution of the number of attempts needed to be made by the man.
(a) if unsuccessful keys are not eliminated from further selection
(b) if they are.

Solution: Let $X$ be the number of attempts needed to open the door.

Case: (a) In this case $X$ can take values $1, 2, 3, \ldots, n$.

So, $P(X = x) = P($success on the first $x-1$ trials and success on the $x$th attempt$)$

\[= \left(\frac{n-1}{n}\right)^{x-1} \cdot \frac{1}{n} \quad \text{[since attempts are made independently]} \]

Case: (b) In this case $X$ can take values $1, 2, 3, \ldots, n$.

So, $P(X = x) = P($success on the first $x-1$ trials and success on the $x$th attempt$)$

\[= \frac{n-1}{n} \cdot \frac{n-2}{n-1} \cdot \frac{n-3}{n-2} \cdots \frac{n-x+1}{n-x+2} \left(\frac{1}{n-x}\right) \]

Ex. 19. There are $n$ tickets in a jar numbered $1, 2, 3, \ldots, n$. Tickets are drawn at random and with replacement from the jar and their numbers are noted; the operation stops as soon as a ticket drawn appears for the second time. Let $X$ be the total number of drawings made, then find its PMF.

Solution: The mass points of $X$ are $2, 3, \ldots, n+1$.

\[P(X = x) = P($the first $x-1$ drawings will give $x-1$ distinct tickets and the $x$th drawing will give a repetition$)\]

\[= \frac{n(n-1)^{x-1} \cdot \frac{x-1}{n}}{n^{x-1}} \]

\[= (x-1) \cdot \left(\frac{n}{x-1}\right) \cdot \left(\frac{x-1}{n}\right) \quad ; \quad x = 2, 3, \ldots, n+1.\]
Ex. 20. The duration (in minutes) of long-distance telephone calls made from a certain city has the distribution function \( F \) given by

\[
F(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 - \frac{1}{2} e^{-x/3} - \frac{1}{2} e^{-[x/2]} & \text{if } x \geq 0
\end{cases}
\]

What is the probability that a telephone call lasts for

(i) more than six minutes?
(ii) less than four minutes?
(iii) exactly three minutes?

What is the conditional probability that the duration of a call is
(iii) less than nine minutes, given that more than five minutes.

\[
\text{Solution:} \\
(i) P(X > 6) = 1 - P(X \leq 6) \\
= 1 - F(6) \\
= e^{-2};
\]

(ii) \( P(X > 4) = F(4) \)

\[
= 1 - \frac{1}{2} (e^{-4/3} + e^{-1});
\]

(iii) \( P(X = 3) = P(X \leq 3) - P(X < 3) \)

\[
= F(3) - F(3-0) = \left(1 - e^{-1}\right) / 2;
\]

(iv) \( P(X < 9 | X > 5) = \frac{P(5 < X < 9)}{P(X > 5)} \)

\[
= \frac{P(X < 9) - P(X \leq 5)}{1 - P(X \leq 5)} = \frac{F(9) - F(5)}{1 - F(5)} = 1 - \frac{e^{-2} + e^{-3}}{e^{-1} + e^{-5/3}};
\]

Remark: - Note that the distribution given above is a mixed distribution, i.e., the weighted average of a discrete and an absolutely continuous distribution. Hence, \( F = \frac{1}{2} F_d + \frac{1}{2} F_c \); therefore,

\[
F_c(x) = \begin{cases} 
1 - e^{-x/3}, & x > 0 \\
0 & \text{otherwise}
\end{cases}
\]

\[F_d(x) = \begin{cases} 
1 - e^{-[x/2]}, & x > 0 \\
0 & \text{otherwise}
\end{cases} \]

This is explained by Decomposition Theorem, see later.
Decomposition Theorem: There may be a distribution whose d.f. is neither discrete nor continuous (absolutely). Such a distribution is called purely singular.

Every distribution function \( F(x) \) can be decomposed into two parts according to \( F(x) = \alpha F_d(x) + (1-\alpha) F_c(x) \), where \( 0 \leq \alpha \leq 1 \), and \( F_d(x), F_c(x) \) are the d.f. of discrete and continuous R.V.s, respectively.

Note that, for \( \alpha = 0 \), \( F(x) = F_c(x) \) is purely continuous.

For \( \alpha = 1 \), \( F(x) = F_d(x) \) is purely discrete. For \( 0 < \alpha < 1 \), then \( F(x) = \alpha F_d(x) + (1-\alpha) F_c(x) \) is neither absolutely continuous nor purely discrete and it is called a mixed distribution.

Further, \( F = \alpha F_d + \beta F_c + \gamma F_S \), where,

\[
F_d \text{ is discrete, } F_c \text{ is absolutely continuous and } F_S \text{ is singular,}
\]

\[
(\alpha > 0, \beta > 0, \gamma > 0, \alpha + \beta + \gamma = 1).
\]

**Ex. 1.** Let for an r.v. \( X \), \( F(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2}, & 0 \leq x < 1 \\ \frac{x - 1}{2}, & 1 \leq x < 2 \\ 1, & 2 \leq x \end{cases} \)

Show that \( F(x) \) can be written as a mixture of two distribution functions.

**Solution:** \( F(x) = \frac{1}{2} (G(x) + H(x)) \), where

\[
G(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x < 1 \\ \frac{x - 1}{2} & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases} \\
H(x) = \begin{cases} 0 & x < 0 \\ \frac{x}{2} & 0 \leq x < 1 \\ 0 & 1 \leq x \end{cases}
\]

Then \( X \) has a mixed distribution.

**Ex. 2.** An r.v. has the d.f. \( F(x) = \begin{cases} 1 - 0.7e^{-x} & \text{if } x \geq 0 \\ 0 & \text{otherwise} \end{cases} \)

Sketch the d.f. Give an e.g. of an r.v. that may be supposed to have the distribution. Also, find

\( a) P[X = 0], \quad b) P[X \leq 4], \quad c) P[3 < X \leq 5] \)

**Solution:**

\[
F(x) \approx \begin{cases} 0 & x < 0 \\ 0.03 & x = 0 \\ 1 & x \geq 0 \end{cases}
\]

Note that, \( F(x) \) has a jump at \( x = 0 \) and \( F(x) \) is discontinuous on \( (0, \infty) \).
Hence, \( F(x) \) is the D.F. of a R.V. \( X \), i.e., neither purely discrete nor purely continuous; i.e., the R.V. \( X \) has an isolated value at \( x=0 \) and can take any value in \((0,1)\) as a continuous R.V.

Write,

\[
F(x) = 0.3 \cdot F_d(x) + (1-0.3) \cdot F_c(x),
\]

where

\[
F_d(x) = \begin{cases} 
0 & \text{if } x<0 \text{ and } x>0 \\
1 & \text{if } x=0
\end{cases}
\]

Now,

(a) \( P[X=0] = P[X \leq 0] - P[X < 0] \)
\[
= F(0) - F(0-0)
\]
\[
= (1-0.72^0)-0
\]
\[
= 0.3
\]

(b) \( P[X \leq 4] = F(4) = 1-0.72^4 \)

(c) \( P[3 < X < 5] = P[X \leq 5] - P[X \leq 3] \)
\[
= F(5) - F(3)
\]
\[
= \{(1-0.72^5)\} - \{(1-0.72^3)\}
\]
\[
= 0.7\left(1-e^5 - e^3\right)
\]

Ex. 3. Let \( X \) be an R.V. with DF

\[
F(x) = \begin{cases} 
0 & \text{if } x<0 \\
\frac{1}{2} & \text{if } x=0 \\
\frac{1}{2} + \frac{x}{2} & \text{if } 0 < x < 1 \\
1 & \text{if } x \geq 1
\end{cases}
\]

Sketch the D.F. Give an example of an R.V. that may be supposed to have the distribution. Also express \( F(x) \) as \( \alpha F_d(x) + (1-\alpha) F_c(x) \).

Solution:

Clearly, \( F(x) \) has a jump at \( x=0 \) and it is continuous on \((0,1)\). So, \( F(x) \) is a D.F. of mixed-type.

Hence, \( F(x) \) is the D.F. of an R.V. which can take the isolated value \( x=0 \) and can take any value between \((0,1)\) as a continuous R.V. We have only one discontinuity point, say \( x=0 \), of \( F(x) \).

Define,

\[
F_d(x) = \begin{cases} 
0 & \text{if } x<0 \\
1 & \text{if } x \geq 0
\end{cases}
\]
\[ F(x) = \alpha F_d(x) + (1-\alpha) F_e(x) \]

When \( \alpha = 0 \), \( F(0) = \alpha F_d(0) + (1-\alpha) F_e(0) \)

\[ \frac{1}{2} = \alpha \cdot 1 + (1-\alpha) \cdot 0 \]

\[ \implies \alpha = \frac{1}{2} \]

Now, for \( 0 < \alpha < 1 \), \( F(x) = \frac{1}{2} F_d(x) + \frac{1}{2} F_e(x) \)

\[ \implies \frac{1}{2} \cdot \frac{x}{2} = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot F_e(x) \]

\[ \implies F_e(x) = \frac{1}{2} \cdot \frac{x}{2} = \frac{1}{4} \cdot \frac{x}{2} \]

For \( x > 1 \), \( F(x) = \frac{1}{2} F_d(x) + \frac{1}{2} F_e(x) \)

\[ \implies 1 = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot F_e(x) \]

\[ \implies F_e(x) = 1 \]

So,

\[ F_d(x) = \begin{cases} 0 & \text{if } \alpha < 0 \text{ and } F_e(x) = \begin{cases} 0 & \text{if } 0 < \alpha < 1 \\ 1 & \text{if } \alpha > 1 \end{cases} \end{cases} \]

with \( \alpha = \frac{1}{2} \).
**EXPECTATION**

**Expectation or Mean:** Let \( X \) be an RV defined on \( (-\infty, \infty, P) \). The expectation on mean of the RV \( X \), denoted by \( E(X) \), is defined by

\[
E(X) = \sum_{i=1}^{\infty} x_i P[X = x_i], \quad \text{provided} \quad \sum_{i=1}^{\infty} |x_i| P[X = x_i] \text{ converges},
\]

if \( X \) is a discrete RV, with mass points \( x_1, x_2, \ldots \),

\[
\int_{-\infty}^{\infty} xf_n(x) \, dx, \quad \text{provided} \quad \int |x| f_n(x) \, dx \text{ converges},
\]

if \( X \) is a continuous RV, with PDF \( f_n(x) \).

**Remark:**

1. Let \( \{x_i, f_i\}; i = 1(1)k; \sum_{i=1}^{k} f_i = n \) be a sample of size \( n \) from the RV \( X \) (discrete). The sample mean is given by

\[
\bar{X} = \frac{1}{n} \sum_{i=1}^{k} x_i f_i \to \sum_{i=1}^{k} x_i P[X = x_i], \quad \text{if} \quad n \to \infty,
\]

by the statistical definition of probability, \( \bar{X} \to E(X) \) as \( n \to \infty \).

2. Consider a discrete random variable \( X \) which takes countable infinite number of values \( x_i \) with positive probabilities \( p_i; i = 1, 2, \ldots \). If \( \sum_{i=1}^{\infty} x_i p_i \) converges conditionally, then the series takes different values \( x_i p_i \) for different re-arrangements of the terms \( x_i p_i \). If the mean \( E(X) \) is to serve as a measure of central tendency of \( X \) we require that the series \( \sum_{i=1}^{\infty} x_i \) is absolutely convergent, i.e., \( \sum_{i=1}^{\infty} |x_i| p_i < \infty \), i.e., \( E(X) \) exists. Note that \( \sum_{i=1}^{\infty} x_i p_i \) may converge but \( \sum_{i=1}^{\infty} |x_i| p_i \) may not, in which case we say that \( E(X) \) does not exist.

**Case-I:** Discrete Distribution with a finite number of mass points

Suppose \( X \) is an RV having a discrete distribution with the PMF \( f \), if the mass points of \( X \) are \( x_1, x_2, \ldots, x_n \). Then by definition, the expected value of \( X \) is \( \sum_{i=1}^{n} x_i f(x_i) \). It is denoted by \( E(X) \) or \( \mu \). Thus \( E(X) \) is the sum of the \( i \)th product of the mass points by their respective probabilities. Ex: Let \( X \) denote the points obtained in throwing a fair die, then \( E(X) = \frac{1}{6} (1+2+3+4+5+6) = 3.5 \).

**Case-II:** Discrete Distribution with countably many mass points

Let \( f \) be the PMF of \( X \) and the mass points be \( x_1, x_2, \ldots \)

we may like to take the sum of the series \( \sum_{i=1}^{\infty} x_i f(x_i) \) as the expected value of \( X \). The \( E(X) \) is said to exist if the series \( \sum_{i=1}^{\infty} x_i f(x_i) \) is absolutely convergent, i.e., \( \sum_{i=1}^{\infty} |x_i| f(x_i) < \infty \). It is defined by

\[
E(X) = \sum_{i=1}^{\infty} x_i f(x_i).
\]
Ex. 1. Let $X$ be the number of trials required to get the first success in a series of Bernoulli Trials with probability of success $p$. Then find the expected value of $X$.

Solution: 

$$f(x) = \begin{cases} p^x (1-p)^{x-1}, & \text{if } x=1, 2, \ldots \frac{1}{p} \\ 0, & \text{otherwise} \end{cases}$$

$$\sum_{i=1}^{\infty} \frac{1}{i} f(x_i) = -p \left[ 1 + 2q + 3q^2 + \ldots \right]$$

The series converges since $0 < q < 1$.

$$E(X) = \frac{p}{(1-q)^2} = \frac{1}{p^2} = \frac{1}{p}$$

Ex. 2. A coin is tossed until a head appears, let $X$ be the number of tosses required. Calculate the value of the number of trials required (including the last toss in which a head has to appear).

Solution: 

$$P(X = i) = \frac{1}{2^i}, \quad i = 1, 2, 3, \ldots$$

$$E(X) = \sum_{i=1}^{\infty} i \cdot \frac{1}{2^i} = \frac{1}{2} + 2 \cdot \frac{1}{2^2} + 3 \cdot \frac{1}{2^3} + \ldots = 2.$$

Indicator Random Variable: - For an event associated with a random experiment we define an RV $I_A(w)$ on each point in the sample space $\Omega$ as:

$$I_A(w) = \begin{cases} 1 & \text{if } w \in A \\ 0 & \text{if } w \notin A \end{cases}$$

$I_A(w)$ is called the indicator RV (function) of the set $A$.

$$\therefore E[I_A(w)] = \sum_{w \in A} P(w) + \sum_{w \notin A} P(w) = \sum_{w \in A} P(w) = P(A).$$

The probability of an event $A$ is the expectation of its indicator RV $I_A(w)$.

Case III: - Absolutely Continuous distribution

Suppose $X$ has the absolutely continuous distribution with pdf $f(x)$. The $E(X)$ is said to exist if the integral $\int_{-\infty}^{\infty} x f(x) dx$ is absolutely convergent, i.e., $\int_{-\infty}^{\infty} |x| f(x) dx < \infty$.

In case, it exists, then expectation is defined as

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$
Ex. 2. Let \( X \) has the pdf \( f(x) = \begin{cases} \frac{1}{6} & 0 < x < 3 \\ 0 & \text{otherwise} \end{cases} \). Find \( E(X) \)?

Solution:

\[
\int_{-\infty}^{\infty} |x| f(x) \, dx = \int_{0}^{3} |x| f(x) \, dx + \int_{0}^{3} |x| f(x) \, dx + \int_{0}^{3} |x| f(x) \, dx
\]

\[
= \int_{0}^{3} |x| f(x) \, dx
\]

\[
= \int_{0}^{3} x \, dx = \left. \frac{x^2}{2} \right|_{0}^{3} = \frac{9}{2}
\]

Hence, the integral is convergent. Hence \( E(X) \) exists and equal to

\[
E(X) = \frac{1}{6} \left[ \frac{9}{2} \right] = \frac{3}{4}.
\]

Ex. 4. Let \( X \) be an RV takes the values \( x_i = (-1)^{i-1} (i+1) \), with probability \( p_i = \frac{1}{i(i+1)} \), \( i \in \mathbb{N} \). Does \( E(X) \) exists?

Solution:

Note that, \( \sum_{i=1}^{\infty} x_i \cdot P[X = x_i] = \sum_{i=1}^{\infty} x_i \cdot p_i = \sum_{i=1}^{\infty} (-1)^{i-1} \cdot (i+1) \cdot \frac{1}{i(i+1)} \)

\[
= \sum_{i=1}^{\infty} (-1)^{i-1} \cdot \frac{1}{i}
\]

Here, \( \sum_{i=1}^{\infty} |x_i| \cdot p_i = \sum_{i=1}^{\infty} |x_i| \cdot p_i = \sum_{i=1}^{\infty} \frac{1}{i} \text{ diverges but}
\]

\( \sum_{i=1}^{\infty} x_i \cdot p_i = \sum_{i=1}^{\infty} (-1)^{i-1} \cdot \frac{1}{i} \text{ converges conditionally.}
\]

As, \( \sum_{i=1}^{\infty} |x_i| \cdot p_i \text{ diverges, } E(X) \text{ doesn't exist; although, } \sum_{i=1}^{\infty} x_i \cdot p_i \text{ converges conditionally.}
\]

Ex. 5. Let \( X \) be an RV with pdf \( f(x) = \begin{cases} \frac{1}{x^2} & 1 < x < \infty \\ 0 & \text{otherwise} \end{cases} \). Show that, \( E(X) \) doesn't exist.

Solution:

\[
\int_{1}^{\infty} |x| f(x) \, dx = \int_{1}^{\infty} x \cdot \frac{1}{x^2} \, dx + \int_{1}^{\infty} x \cdot \frac{1}{x^2} \, dx = \int_{1}^{\infty} \frac{1}{x} \, dx = \infty.
\]

\( \therefore E(X) \) does not exist.
Ex. 6. X has a continuous distribution pdf \( f(x) = \frac{1}{\pi(1+x^2)}, x \in \mathbb{R} \). Show that \( E(X) \) does not exist.

Solution:
Note that \( \int_{-\infty}^{\infty} |x| f(x) \, dx = \int_{0}^{\infty} x \cdot \frac{1}{\pi(1+x^2)} \, dx = \frac{1}{\pi} \int_{0}^{\infty} \frac{x \, dx}{1+x^2} = \frac{1}{2\pi} \int_{0}^{\infty} \frac{dt}{t} \left[ \log t \right]_{0}^{\infty} = \infty. \)

\( \therefore E(X) \) does not exist.

Ex. 7. A target is made of 3 concentric circles of radii \( \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}} \) feet. Shots within the inner circle count 4 points, within the next ring 3 points and within the third ring 2 points. Shots outside the target count zero. Let \( X \) be the distance of the hit from the centre (in feet) and the pdf of \( X \) be

\[ f(x) = \begin{cases} \frac{2}{\pi(1+x^2)}, & x > 0 \\ 0, & \text{otherwise} \end{cases} \]

What will be the expected value of the score (a) one shot; (b) a set of 5 shots?

Solution:
\( X \): the hit from centre, \( Y \): the score.

<table>
<thead>
<tr>
<th>Mass points of ( Y )</th>
<th>Corresponding Probabilities</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>( P(Y = 4) = P(X \in (0, \frac{1}{\sqrt{3}})) )</td>
</tr>
<tr>
<td>3</td>
<td>( P(Y = 3) = P(X \in (\frac{1}{\sqrt{3}}, 1)) )</td>
</tr>
<tr>
<td>2</td>
<td>( P(Y = 2) = P(X \in (1, \frac{1}{\sqrt{3}})) )</td>
</tr>
<tr>
<td>0</td>
<td>( P(Y = 0) = P(X \in (\frac{1}{\sqrt{3}}, \infty)) )</td>
</tr>
</tbody>
</table>

\( a) \quad E(Y) = 4 \cdot P(Y = 4) + 3 \cdot P(Y = 3) + 2 \cdot P(Y = 2) + 0 \cdot P(Y = 0) \)
\[ = 4 \cdot P \left[ 0 < X \leq \frac{1}{\sqrt{3}} \right] + 3 \cdot P \left[ \frac{1}{\sqrt{3}} < X \leq 1 \right] + 2 \cdot P \left[ 1 < X \leq \frac{1}{\sqrt{3}} \right] \]

Now, \( P[a < X \leq b] = F(b) - F(a) = \int_{a}^{b} f(x) \, dx = \int_{a}^{b} \frac{2}{\pi(1+x^2)} \, dx \)
\[ = \frac{2}{\pi} \left[ \tan^{-1}x \right]_{a}^{b} = \frac{2}{\pi} \left[ \tan^{-1}b - \tan^{-1}a \right] \]

So, \( E(Y) = \frac{2}{\pi} \left\{ 4 \left( \frac{\pi}{6} - 0 \right) + 3 \left( \frac{\pi}{4} - \frac{\pi}{6} \right) + 2 \left( \frac{\pi}{3} - \frac{\pi}{4} \right) \right\} = \frac{43}{6} \).

\( b) \quad \text{Let } Y_i \text{ denote the score in the } i^{th} \text{ shot}, i = 1(1)5. \text{ Then, total score in 5 shots} \)

\[ Z = \sum_{i=1}^{5} E(Y_i) = 5 \cdot E(Y_1) = 5 \cdot \frac{43}{6} = \frac{215}{6} \]
Theorems on Expectation:

Theorem 1. If the RV $X = c$, a finite real number with probability 1, i.e. $X = c$ almost everywhere, then $E(X) = c$.

Proof: 
\[ E(X) = \int X \, dP = \int_{x=c} x \, dP = c \cdot P(X = c) = c. \]

Theorem 2. If $c$ is a finite real number and $E(X)$ exists then $E(cX)$ also exists and equals to $c \cdot E(X)$.

Proof: Since $E(X)$ exists, so $\int |X| \, dP < \infty$.

Now $\int |cX| \, dP = |c| \int |X| \, dP$, since $c$ is finite, so $\int |cX| \, dP < \infty$.

So, $E(cX)$ exists and equals $c \int X \, dP = cE(X)$.

Theorem 3. If $X$ and $Y$ are both RV's, and $E(X)$ and $E(Y)$ exist, then $E(X + Y)$ exists and equals to $E(X) + E(Y)$.

Solution: 
$E(X)$ & $E(Y)$ exist, so $\int |X| \, dP < \infty$, $\int |Y| \, dP < \infty$.

Now, 
\[ \int |X + Y| \, dP = \int |X| \, dP + \int |Y| \, dP < \infty. \]

So, $E(X + Y)$ exists. 
\[ E(X + Y) = \int X \, dP + \int Y \, dP = E(X) + E(Y). \]

Theorem 4. If $E(X)$ exists, then $|E(X)| \leq E(|X|)$.

Proof: 
\[ |\int X \, dP| \leq \int |X| \, dP \Rightarrow |E(X)| \leq E(|X|), \text{ since } E(X) \text{ exists,} \]

so $X$ is integrable w.r.t. $P$.

Theorem 5. If $E(X)$ exists and $a$ and $b$ are real numbers $\exists a \leq X \leq b$, then $a \leq E(X) \leq b$.

Solution: 
$a \leq X \leq b$
\[ \int a \, dP \leq \int X \, dP \leq \int b \, dP \]

\[ \therefore a \leq E(X) \leq b. \]
Problem: 2. The RV \( X \) takes non-negative integer values. Show that 
\[
E(X) = \sum_{k=0}^{\infty} P(X > k),
\]
provided the series on the right hand side converges.

Solution: 
\[
E(X) = \sum_{\alpha=0}^{\infty} \alpha P[\alpha] \\
= P(1) + 2P(2) + 3P(3) + \cdots \\
= \left\{ P(1) + P(2) + P(3) + \cdots \right\} \\
+ \left\{ P(2) + P(3) + \cdots \right\} \\
[\text{By the rearrangement of the terms}]
= P[X > 0] + P[X > 1] + \cdots \\
= \sum_{k=0}^{\infty} P(X > k).
\]

Problem: 3. Find the mean of the truncated Poisson distribution with pmf 
\[
f(\alpha) = \begin{cases} \frac{\lambda^{\alpha} e^{-\lambda}}{\alpha!} & \text{if } \alpha = 1, 2, 3, \ldots \\ 0 & \text{otherwise} \end{cases}
\]

Solution: 
\[
E(X) = \sum_{\alpha=1}^{\infty} \alpha \frac{\lambda^{\alpha} e^{-\lambda}}{\alpha!} \\
= \frac{\lambda}{1-e^{-\lambda}} \sum_{\alpha=1}^{\infty} \frac{\lambda^{\alpha} e^{-\lambda}}{(\alpha-1)!} \\
= \frac{\lambda}{1-e^{-\lambda}}.
\]

Problem: 4. Show that for triangular distribution with pdf 
\[
f(\alpha) = \begin{cases} \frac{\alpha}{\alpha} \left[ 1 - \frac{|\alpha - \Theta|}{\alpha} \right] & \text{if } |\alpha - \Theta| \leq \alpha \\ 0 & \text{otherwise} \end{cases}
\]
the mean is equal to \( \Theta \).

Solution: 
\[
E(X) = E(X - \Theta + \Theta) = E(X - \Theta) + E(\Theta) \\
= \int_{0}^{\alpha} (\alpha - \Theta) f(\alpha) d\alpha + \Theta \\
= \Theta + \int_{-\alpha}^{0} (\alpha - \Theta) \left[ 1 + \frac{\alpha - \Theta}{\alpha} \right] d\alpha + \int_{0}^{\alpha} (\alpha - \Theta) \left[ 1 - \frac{\alpha - \Theta}{\alpha} \right] d\alpha \\
= \Theta + \int_{0}^{\alpha} 2d\alpha + \int_{0}^{\alpha} 2d\alpha + \int_{0}^{\alpha} 2d\alpha - \int_{0}^{\alpha} 2d\alpha \\
= \Theta.
\]
Expectation of a function of Random Variable: - If \( X \) is an RV defined on \((-\infty, a, P)\) and \( g(X) \) be a function of \( X \), then \( g(X) \) is also an RV defined on \((-\infty, a, P)\) and the expected value of \( g(X) \) is defined as

\[
E[g(X)] = \begin{cases} 
\sum_{i=1}^{\infty} g(x)P[X = x], & \text{provided } \sum g(x)P[X = x] \text{ converges, if } X \text{ is of discrete type.} \\
\int g(x)f_X(x)dx, & \text{provided } \int |g(x)|f_X(x)dx \text{ converges, if } X \text{ is of continuous type with PDF } f_X(x). 
\end{cases}
\]

Alternative Definition: - Let \( g(X) \) be a function of \( X \) which is itself an RV, then expectation of \( g(X) \) is said to exist if 

\[
\int_{-\infty}^{\infty} |g(x)|dF(x) < \infty.
\]

Then 

\[
E(g(X)) = \int_{-\infty}^{\infty} g(x)dF(x).
\]

Problem: - 5. From an urn with \( a \) white and \( b \) black balls, \( c \) balls are taken at random and transferred to another urn which contains \( x \) white and \( y \) black balls. Show that the probability of getting a white ball from the 2nd urn after the transfer is 

\[
\frac{ca}{\alpha + \beta + c}.
\]

Solution: - \( X \) : No. of white balls transferred to the 2nd urn.

Let \( A \) : Event that a white ball from 2nd urn is obtained.

\( B_x \) : Event that \( x \) white balls are transferred, \( x = 0, 1, 2, \ldots \).

From the theorem of total probability,

\[
P(A) = \sum_{x} P(B_x)P(A|B_x) = \sum_{x} \frac{\alpha \cdot \beta}{\alpha + \beta + c}.
\]

\[
= \frac{\alpha}{\alpha + \beta + c} \sum_{x} P(B_x) + \sum_{x} \frac{\alpha \cdot \beta}{\alpha + \beta + c} P(B_x)
\]

\[
= \frac{\alpha}{\alpha + \beta + c} \times 1 + \frac{E(X)}{\alpha + \beta + c}
\]

\[
= \frac{\alpha}{\alpha + \beta + c} + \frac{ca}{\alpha + \beta + c}.
\]

\[
= \frac{ca}{\alpha + \beta + c} + \frac{\alpha}{\alpha + \beta + c}
\]

\[
= \frac{ca + \alpha}{\alpha + \beta + c}.
\]

\[
P(B_x) = \frac{(\alpha)(c-x)}{(\alpha + b)^c}.
\]

\(x = 0, 1, 2, \ldots \).

\[
E(X) = \sum_{x} xP(B_x) = \frac{ca}{\alpha + b}.
\]

Mean of Hypergeometric Distribution.
Result 1. If \( X \) is a non-negative integer valued random variable, then show that \( E(X) = \sum_{x=0}^{\infty} \{1 - F(x)\} \), provided \( E(X) \) exists.

Proof: \( E(X) = \sum_{x=0}^{\infty} x \cdot p_x \), where \( p_x = P[X=x] \):

\[
= 0 \cdot p_0 + 1 \cdot p_1 + 2 \cdot p_2 + \ldots \ldots \ldots \ldots \\
= (p_1 + p_2 + \ldots) + (p_2 + p_3 + \ldots) + \ldots \\
= \sum_{x=0}^{\infty} [1 - P[X \leq x]] = \sum_{x=0}^{\infty} \{1 - F(x)\} \\
\]

Result 2. Suppose \( X \) is a non-negative R.V whose mean exists and equals \( \mu \). Prove that

(a) \( \lim_{x \to \infty} x[1 - F(x)] = 0 \),

and hence in case \( X \) is absolutely continuous, then

(b) \( \int_0^{\infty} [1 - F(x)]dx = \mu \).

Proof: (a) \( E(X) \) exists and equals to \( \mu \). Note that

\[
\lim_{x \to \infty} \int_0^x u \cdot dF(u) = 0. \\
\]

Again, \( \int_0^x u \cdot dF(u) \geq x \cdot \int_0^x dF(u) = x \cdot [1 - F(x)] \geq 0 \) \( \because X \) is non-negative

\[
\implies \int_0^\infty u \cdot dF(u) \geq \lim_{x \to \infty} x \cdot [1 - F(x)] \geq 0. \\
\]

\[
\therefore \lim_{x \to \infty} x \cdot [1 - F(x)] = 0. \\
\]

\[
\cdot \lim_{x \to \infty} x \cdot [1 - F(x)] = 0. \\
\]
(b) \[ \mu = \int_0^\infty x f(x) \, dx \]
\[ = \int_0^\infty - \alpha [-d(1-F(x))] \, dx \quad \text{since } f(x) = -\frac{d}{dx} [1-F(x)] \]
\[ = \left[ -\alpha (1-F(x)) \right]_0^\infty + \int_0^\infty [-1-F(x)] \, dx \]
\[ = \lim_{u \to \infty} \left[ -\alpha (1-F(x)) \right]_0^u + \int_0^\infty [-1-F(x)] \, dx \]
\[ = \lim_{u \to \infty} \left[ -u (1-F(u)) \right] + \int_0^\infty [-1-F(x)] \, dx \]
\[ = \int_0^\infty [-1-F(x)] \, dx \quad \text{[From part (a)]} \]

Also, see:

\[ E(x^2) = \int_0^\infty x^2 f(x) \, dx \]
\[ = \int_0^\infty x^2 [-d(1-F(x))] \, dx \]
\[ = \left[ -\alpha^2 (1-F(x)) \right]_0^\infty + 2\int_0^\infty x [1-F(x)] \, dx \]
\[ = 2\alpha \int_0^\infty [1-F(x)] \, dx . \]

**Remark:**
1. \[ E(X) = \int_0^\infty [1-F(x)] \, dx . \]

![Graph of F(x)](image)

Geometry of \( E(X) := \)

2. \[ E(X) = \int_0^\infty x f(x) \, dx = \int_0^t x f(x) \, dx + \int_t^\infty x f(x) \, dx \]

Hence \( E(X) \) exists if and only if \( \int_0^\infty x f(x) \, dx \) converges.

Again, \[ E(X) = \int_0^\infty [1-F(x)] \, dx = \int_0^t [1-F(x)] \, dx + \int_t^\infty [1-F(x)] \, dx \]

\( E(X) \) exists if and only if \( \int_t^\infty [1-F(x)] \, dx \) converges.

...
Result: 3. (a) If $X$ is any RV then $E(X)$ exists, then show that

$$E(X) = \int_0^\infty [1 - F(x) - F(-x)] \, dx,$$

(b) Hence show that $E(X^2) = \int_0^\infty 2x [1 - F(x) - F(-x)] \, dx$.

Proof:

(a) $\int_0^\infty [1 - F(x) - F(-x)] \, dx$

$$= \int_0^\infty [1 - F(x)] \, dx - \int_0^\infty F(-x) \, dx$$

$$= \int_0^\infty x f(x) \, dx - \int_0^\infty (-x) f(x) \, dx \quad \text{[From Result 2(b)]}$$

$$= \int_0^\infty x f(x) \, dx + \int_\infty^0 x f(x) \, dx$$

$$= \int_0^\infty x f(x) \, dx = E(X).$$

(b) $\int_0^\infty 2x [1 - F(x) - F(-x)] \, dx$

$$= \int_0^\infty 2x [1 - F(x)] \, dx - \int_0^\infty 2x F(-x) \, dx$$

$$= \int_0^\infty x^2 f(x) \, dx + \int_0^\infty x^2 f(x) \, dx \quad \text{[From Result 2(b)]}$$

$$= \int_0^\infty x^2 f(x) \, dx$$

$$= E(X^2).$$
Ex. 1. Evaluate $E(X)$ for the RV $X$ with DF

$$F(x) = \begin{cases} 
  0 & \text{if } x < 0 \\
  1 - (1-x)^n & \text{if } 0 \leq x < 1 \\
  1 & \text{if } x \geq 1 
\end{cases}$$

Solution: For a non-negative continuous RV $X$ with DF $F(x)$,

$$E(X) = \int_0^\infty [1 - F(x)] \, dx$$

$$= \int_0^1 [1 - F(x)] \, dx + \int_1^\infty [1 - F(x)] \, dx$$

$$= \int_0^1 (1-x)^n \, dx + \int_1^\infty 0 \, dx$$

$$= \int_0^1 (1-x)^n \, dx = \left[ -\frac{(1-x)^{n+1}}{n+1} \right]_0^1 = \frac{1}{n+1}.$$

Ex. 2. Let $X$ and $Y$ be two non-negative continuous RVs with respective DFs $F(x)$ and $G(x)$ ($F(x) \geq G(x)$ $\forall x \geq 0$). If $E(X)$ and $E(Y)$ exist, then show that $E(X) \leq E(Y)$.

Solution: As $X$ and $Y$ are two non-negative continuous RVs, then

$$E(X) = \int_0^\infty [1 - F(x)] \, dx, \quad E(Y) = \int_0^\infty [1 - G(x)] \, dx.$$ 

Hence,

$$F(x) \geq G(x) \quad \forall x \geq 0$$

$$1 - F(x) \leq 1 - G(x) \quad \forall x \geq 0$$

$$\Rightarrow \int_0^\infty [1 - F(x)] \, dx \leq \int_0^\infty [1 - G(x)] \, dx, \quad \forall x \geq 0$$

$$\Rightarrow E(X) \leq E(Y)$$

- Geometric significance of $G(x)$ and $F(x)$:

\[ Y = F(x) \rightarrow \quad Y = G(x) \]
Ex. 3. Find the expected number of throws of a fair dice until a six is obtained.

**Solution:** Consider 'getting a six' in a throw of a fair dice as success.

Let, \( X \) be the number of throws required to get first success.

Then \( P[X = x] = P[\text{The first } (x-1) \text{ throws result in failures and a success occurs, at the } x^{th} \text{ throw}] \)

\[
= \frac{1}{6} \cdot \left( \frac{5}{6} \right)^{x-1}, \quad x = 1, 2, 3, \ldots
\]

\[
E(X) = \sum_{x=1}^{\infty} x \cdot P[X = x] = \sum_{x=1}^{\infty} x \cdot \left( \frac{5}{6} \right)^{x-1} \cdot \left( \frac{1}{6} \right)
\]

\[
= \frac{1}{6} \sum_{x=1}^{\infty} x \cdot \left( \frac{5}{6} \right)^{x-1}
\]

\[
= \frac{1}{6} \cdot \left( 1 - \frac{5}{6} \right)^{-2} = 6.
\]

Ex. 4. Balls are taken one by one with replacement out of an urn containing \( a \) white balls and \( b \) black balls until the first white ball is drawn. What is the expectation of the number of black balls proceeding the first white ball?

**Solution:**

\[
P[ \text{A white ball is drawn}] = \frac{a}{a+b} = p,
\]

\[
P[ \text{A black ball is drawn}] = \frac{b}{a+b} = q.
\]

\( X \): An RV denoting the number of black balls drawn.

\[
P[\text{No. of black balls drawn preceding the first white ball}] = pq^\alpha; \alpha = 0, 1, 2, \ldots
\]

So,

\[
\sum_{\alpha=0}^{\infty} pq^\alpha = 1
\]

Now,

\[
\sum_{\alpha=0}^{\infty} p(1-p)^\alpha = 1
\]

Differentiating \( \sum \) \( \alpha \) \( cont. \),

\[
\frac{d}{dx} \sum_{\alpha=0}^{\infty} pq^\alpha = \frac{d}{dx} \sum_{\alpha=0}^{\infty} \alpha pq^\alpha = 0
\]

\[
\Rightarrow \sum_{\alpha=0}^{\infty} \alpha pq^\alpha = \frac{1}{1-q}
\]

\[
\Rightarrow \frac{1}{q} \sum_{\alpha=0}^{\infty} \alpha p q^\alpha = \frac{1}{p}
\]

\[
\Rightarrow E(X) = \frac{q}{1-q} = \frac{\frac{b}{a+b}}{\frac{a}{a+b}} = \frac{a}{a+b}\left[ \because a+b \neq 0 \right]
\]
Ex. 5. From an urn containing \( N \) identical tickets numbered 1 to \( N \), \( n \) tickets are drawn without replacement, let \( X \) be the largest number drawn. Hence, find \( E(X) \). Also, show that, for large \( N \),

\[
E(X) \approx \frac{nN}{n+1}.
\]

**Solution:** Note that, \( P[X \leq \alpha] = \) Probability that the largest number in \( n \) drawn tickets with replacement is less than or equal to \( \alpha \).

\[
= \text{Prob. that each of } n \text{ drawn ticket is less than or equal to } \alpha,
\]

\[
\Rightarrow P[X \leq \alpha] = \frac{\alpha^n}{N^n}, \quad \alpha = 1, 2, \ldots, N.
\]

Hence, \( F_X(\alpha) = \frac{\alpha^n}{N^n}, \quad \alpha = 1, 2, \ldots, N \).

For non-negative integer valued R.V. \( X \), we have,

\[
E(X) = \sum_{\alpha=1}^{\infty} [1 - F_X(\alpha)]
\]

\[
= \sum_{\alpha=1}^{N-1} \left( 1 - \frac{\alpha^n}{N^n} \right) + \sum_{\alpha=N}^{\infty} (1 - 1)
\]

\[
= N - \frac{1}{N^n} \sum_{\alpha=0}^{N-1} \alpha^n.
\]

\( \text{For large } N, \)

\[
\frac{1}{N} \sum_{\alpha=0}^{N-1} \left( \frac{\alpha}{N} \right)^n \approx \int_0^1 y^n \, dy, \quad \text{for large } N,
\]

\[
= \frac{1}{n+1}.
\]

\( \Rightarrow \frac{1}{N} \sum_{\alpha=0}^{N-1} \alpha^n = \frac{N}{n+1}, \quad \text{for large } N, \)

\( \text{Now, } E(X) = N - \frac{1}{N^n} \sum_{\alpha=0}^{N-1} \alpha^n \)

\[
= N - \frac{N}{n+1}, \quad \text{for large } N,
\]

\[
= \frac{nN}{n+1}, \quad \text{for large } N.
\]
Moments: ~

If we take \( g(x) = x^n \), then \( \mu_n = E(x^n) \) if it exists, is called the \( n \)th order raw moment of \( X \).

For a sample data, \( \{ (x_i, f_i) : i = 1 \} K, \sum_{i=1}^{K} f_i = n \} \), the \( n \)th order sample raw moment, \( m_n = \frac{1}{n} \sum_{i=1}^{n} x_i^n \).

As \( n \to \infty \), by the definition of probability,

\[
E(x^n) = \lim_{n \to \infty} m_n
\]

Here, \( \mu_1 = E(X) \) is the mean of \( X \).

Now, take \( g(x) = [x - E(X)]^n \), then

\( \mu_n = E[(X - E(X))^n] \) if it exists, is called the \( n \)th order central moment of \( X \).

Variance of \( X \): The 2nd order central moment of \( X \),

\( \mu_2 = E[(X - E(X))^2] \), is called the variance of \( X \) and it is denoted by \( \text{Var}(X) \) or \( \sigma_X^2 \).

Now,

\[
\sigma_X^2 = \text{Var}(X) = E\{X - E(X)\}^2
\]

\[
= E((X - \mu)^2), \quad \mu = E(X)
\]

\[
= E(X^2 - 2\mu X + \mu^2)
\]

\[
= E(X^2) - 2\mu E(X) + \mu^2
\]

\[
= E(X^2) - \mu^2
\]

\[
= E(X^2) - E^2(X).
\]

- **Ex.1.** If \( X \) is a non-negative R.V., then \( E(X) \geq 0 \).

  **Sol.** Let \( X \) be a discrete R.V. with mass points \( \alpha_1, \alpha_2, \ldots \).

  As \( X \) is non-negative R.V., \( \alpha_i \geq 0 \) \( \forall i = 1, 2, \ldots \).

  \[ E(X) = \sum_{i=1}^{\alpha_i} \alpha_i P[X = \alpha_i] \geq 0 \] as \( P[X = \alpha_i] > 0 \) \( \forall i \).

- **Ex.2.** If \( X \) is a non-negative R.V. and \( E(X) = 0 \), then \( P[X = 0] = 1 \).

  **Sol.** \( E(X) = 0 = \sum_{i=1}^{\alpha_i} \alpha_i P[X = \alpha_i] \), where \( \alpha_i \geq 0 \), \( P[X = \alpha_i] > 0 \) \( \forall i \).

  So, \( \alpha_i \cdot P[X = \alpha_i] = 0 \) \( \forall i \).

  \( \Rightarrow \alpha_i = 0 \) \( \forall i \) as \( P[X = \alpha_i] > 0 \).

  \( \Rightarrow X = 0 \) with probability 1.

  i.e. \( P[X = 0] = 1 \).
Problem: Show that if $X$ is an RV such that $P[a \leq X \leq b] = 1$, then $E(X)$ and $\text{Var}(X)$ exist and $a \leq E(X) \leq b$ and $\text{Var}(X) \leq \frac{(b-a)^2}{4}$.

Solution:
Let $X$ be an RV of discrete type.

Then $a \leq X_i \leq b$ for all $i$, so $P[a \leq X \leq b] = 1$.

$\Rightarrow \quad a P[X = X_i] \leq \alpha; P[X = X_i] \leq b P[X = X_i]$

$\Rightarrow \quad a \leq E(X) \leq b$.

Now, $\text{Var}(X) = E[(X-E(X))^2] \leq E\left(X - \frac{a+b}{2}\right)^2$.

[$*: \text{Variance is least mean square deviation}$]

Note,

$a \leq X_i \leq b$

$a - \frac{a+b}{2} \leq X_i - \frac{a+b}{2} \leq b - \frac{a+b}{2}$

$\Rightarrow \quad \left(X_i - \frac{a+b}{2}\right)^2 \leq \left(b - \frac{a+b}{2}\right)^2$

$\Rightarrow \quad \sum_{i=1}^{n} \left(X_i - \frac{a+b}{2}\right)^2 P[X = X_i] \leq \left(b - \frac{a+b}{2}\right)^2 \sum_{i=1}^{n} P[X = X_i]$

$\Rightarrow \quad E\left(X - \frac{a+b}{2}\right)^2 \leq \left(b - \frac{a+b}{2}\right)^2 \Rightarrow \text{Var}(X) \leq \frac{(b-a)^2}{4}$.

Result: 1. Show that $E(X - a)^2$ is minimum when $a = E(X)$; $\text{Var}(X) \leq E(X-a)^2$.

Solution:

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

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

$= \text{Var}(X) + (E(X)-a)^2$

i.e., $E(X-a)^2 \geq \text{Var}(X)$, since $(E(X)-a)^2 \geq 0$.

"=" holds when $E(X) = a$. So, $E(X-a)^2$ is minimum when $E(X) = a$.

Result: 2. Suppose that for the variable $X$, the 2nd order moment exists, and $\mu_2'(A) > \mu_2$.

[Standard deviation is the least RMS deviation]

Proof:

$\mu_2'(A) = E(X-A)^2$

$= E[(X-\mu) + (\mu - A)]^2$

$= E(X-\mu)^2 + (\mu - A)^2 + 2(\mu - A) E(X-\mu)$

$= \mu_2 + (\mu - A)^2$ ["$\because E(X-\mu) = 0$]

$\Rightarrow \quad \mu_2'(A) > \mu_2$.

"=" sign holds if $A = \mu$. 

Scanned by CamScanner
Ctanujit Classes
Page No. 93
Theorem: If the moment of order $p$ exists for an R.V. $X$, then the moment of order $q$ ($q < p$) also exists.

Proof: Let $X$ be a continuous R.V. with PDF $f(x)$.

Note that

\[
\int_{-\infty}^{\infty} |x|^q f(x) \, dx
\]

\[
= \int_{|x| \leq 1} |x|^q f(x) \, dx + \int_{|x| > 1} |x|^q f(x) \, dx
\]

\[
\leq \int_{|x| \leq 1} f(x) \, dx + \int_{|x| > 1} |x|^q f(x) \, dx
\]

\[
\leq P[|X| \leq 1] + \int_{|x| > 1} |x|^q f(x) \, dx
\]

[For $|x| \leq 1$, $|x|^q \leq 1$, and for $|x| > 1$, $|x|^q < |x|^p$ as $q < p$]

\[
\int_{-\infty}^{\infty} |x|^q f(x) \, dx \leq P[|X| \leq 1] + \int_{|x| > 1} |x|^p f(x) \, dx
\]

\[
\leq 1 + \int_{|x| > 1} |x|^p f(x) \, dx < \infty, \text{ as } E(X^p) \text{ exists.}
\]

Hence, $E(X^q)$ exists, provided $E(X^p)$ exists, for $q < p$.

Ex. 1. Give an example of discrete distribution whose mean exists but variance does not.

Sol.: Let $X$ be a discrete R.V. with pmf

\[
P[X = i] = \begin{cases} k \cdot \frac{1}{i^3}, & i = 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}
\]

cohere, $k = \frac{1}{\sum_{i=1}^{\infty} \frac{1}{i^3}}$

Note that,

\[
\sum_{i=1}^{\infty} |i| P[X = i] = \sum_{i=1}^{\infty} i \cdot \frac{k}{i^3} = k \sum_{i=1}^{\infty} \frac{1}{i^2}, \text{ converges,}
\]

But,

\[
\sum_{i=1}^{\infty} i^2 P[X = i] = \sum_{i=1}^{\infty} i^2 \cdot \frac{k}{i^3} = k \sum_{i=1}^{\infty} \frac{1}{i}, \text{ diverges.}
\]

Hence, $E(X)$ exists but $E(X^2)$ does not.

$\Rightarrow$ $E(X)$ exists but $\text{Var}(X)$ does not exist.
Ex. 2. Give an example of a continuous distribution where mean exists but variance does not.

Sol.: - Let \( X \) be a continuous RV with PDF

\[
    f(x) = \begin{cases} \frac{2}{x^3}, & x > 1 \\ 0, & \text{otherwise} \end{cases}
\]

Note that,

\[
    \int_{-\infty}^{\infty} |x| f(x) \, dx = \int_{1}^{\infty} x \cdot \frac{2}{x^3} \, dx = \lim_{t \to \infty} \int_{1}^{t} \frac{2}{x^2} \, dx
\]

\[
    = \lim_{t \to \infty} \left[ -\frac{1}{x} \right]_1^t
\]

\[
    = \lim_{t \to \infty} \left( 1 - \frac{1}{t} \right)
\]

But,

\[
    \int_{-\infty}^{\infty} x^2 f(x) \, dx = \int_{1}^{\infty} x^2 \cdot \frac{2}{x^3} \, dx = \int_{1}^{\infty} \frac{2}{x} \, dx = \lim_{t \to \infty} \int_{1}^{t} \frac{2}{x} \, dx
\]

\[
    = \lim_{t \to \infty} \left[ \ln x \right]_1^t
\]

\[
    = \lim_{t \to \infty} \ln t
\]

Hence, \( E(X) \) exists but \( E(X^2) \) or \( \text{Var}(X) \) does not.

Remark: - Consider the PMF

\[
    P[X = k] = \begin{cases} \frac{k}{x^{x+2}}, & x = 1, 2, 3, \ldots \\ 0, & \text{otherwise} \end{cases}
\]

Note that \( E(X^k) \) exists, but \( E(X^{k+1}) \) does not.

2. Consider the PDF

\[
    f(x) = \begin{cases} \frac{x^{x+1}}{x^{x+2}}, & x > 1 \\ 0, & \text{otherwise} \end{cases}
\]

Note that \( \mu \) exists but \( \mu^{n+1} \) does not.
Problem 1. Assume that n random variables \( X_1, X_2, \ldots, X_n \) are independent and each takes the values 1 and -1 with probability \( p \) and \( 1-p \), respectively. Find the expectation and variance of the product of the random variables.

Solution:

Let \( X = X_1 X_2 \cdots X_n \).

\[ E(X) = E(X_1) E(X_2) \cdots E(X_n) \quad \text{[Due to independence of } X_i's]\]

\[ E(X) = \prod_{i=1}^{n} E(X_i) \]

\[ = \prod_{i=1}^{n} \left\{ (1)p + (-1)(1-p) \right\} \]

\[ = (2p-1)^n. \]

\[ \text{Var}(X) = E(X^2) - E^2(X) \]

\[ = E(X_1^2 X_2^2 \cdots X_n^2) - (2p-1)^{2n} \]

\[ = \prod_{i=1}^{n} E(X_i^2) - (2p-1)^{2n}, \text{ by product law of} \]

\[ = \prod_{i=1}^{n} \left\{ (1)^2 p + (-1)^2 (1-p) \right\} - (2p-1)^{2n} \]

\[ = 1 - (2p-1)^{2n}. \]

Problem 2. If \( X \) is a discrete R.V. and \( E(X^2) = 0 \), show that \( P(X = 0) = 1 \).

Deduce that if \( \text{Var}(X) = 0 \) then \( P(X = \mu) = 1 \), where \( \mu = E(X) \).

Solution:

(i) \[ E(X^2) = 0 \]

\[ \Rightarrow \sum_x x^2 P(x) = 0 \]

\[ \Rightarrow 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2^2 \cdot P(X = 2) + \cdots = 0 \quad \text{(4)} \]

From (4), we get \[ P(X = j) = 0 \quad \forall \quad j = 1, 2, 3, \ldots. \]

\[ \sum_x P(x) = 1 \Rightarrow P(X = 0) = 1. \]

(ii) \[ E(X) = \mu \]

\[ \text{Var}(X) = 0 \]

\[ \Rightarrow E(X^2) - E^2(X) = 0 \]

\[ \Rightarrow E(X^2) = \mu^2. \]

\[ \Rightarrow \sum_x x^2 P(x = x) = \mu^2 \]

\[ \Rightarrow 0 \cdot P(X = 0) + 1 \cdot P(X = 1) + 2^2 \cdot P(X = 2) + \cdots + \mu^2 P(X = \mu) + \cdots \]

\[ \Rightarrow \mu^2 P(X = \mu) = \mu^2, \text{where } P(x = j) = 0 \forall j \neq \mu, j > 0. \]

\[ \Rightarrow P(X = \mu) = 1. \]
Problem 3. Consider the distribution of an R.V. X with pdf
\[ f(x) = \begin{cases} \alpha^\beta x^{\beta} & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \]
coleurs, both \( \alpha \) and \( \beta \) are positive, show that the moment of order \( n \) exists if \( n < \beta \). Assuming \( \beta > 2 \), find the mean and variance of the distribution.

Solution:
\[ E(X^n) = \int_0^\infty x^n \frac{\beta x^{\beta - 1}}{\alpha x^{\beta + 1}} \, dx = \int_0^\infty \frac{\beta x^{\beta - 1}}{\alpha} x^{n-1} \, dx \]
\[ = \frac{\beta x^{\beta}}{\alpha} \left[ \frac{x^{n-\beta}}{n-\beta} \right]_0^\infty \]
\[ = \frac{\beta x^{\beta}}{n-\beta} \left( 1 - x^{n-\beta} \right), \text{ when } n < \beta. \]
\[ = \frac{\beta x^{\beta}}{n-\beta}, \quad \text{so, } E(X) = \frac{\beta x^{\beta}}{\beta - 1}; \quad E(X^2) = \frac{\alpha^2 \beta^2}{\beta - 2}. \]
\[ \text{Var}(X) = E(X^2) - E^2(X) = \frac{\alpha^2 \beta^2}{\beta - 2} - \frac{\beta^2 \alpha^2}{(\beta - 1)^2}. \]

Problem 4. Let \( X \) be an R.V. with PMF \( p(x) = c \cdot \left( \frac{2N-x}{N} \right) x \), \( x = 0, 1, \ldots, N \), otherwise.

(i) Find the constant \( c \); (ii) Find \( p(x+1) \) and \( E(X) \).

Solution:
(i) \( \sum_x p(x) = 1 \Rightarrow \sum_{x=0}^N c \left( \frac{2N-x}{N} \right) 2^x = 1 \)
\[ \Rightarrow c \left[ \binom{2N}{N} + \binom{2N-1}{N} 2 + \cdots \right] = 1. \]

(ii) \[ \frac{p(x+1)}{p(x)} = \frac{\binom{2N-(x+1)}{N}}{\binom{2N-x}{N}} 2 = \frac{2(N-x)}{(2N-x)} \]
\[ \Rightarrow \sum_{x=0}^N \frac{(2N-x) p(x+1)}{p(x)} = 2 \sum_{x=0}^N (N-x) p(x) \]
\[ \Rightarrow \sum_{x=0}^N \frac{2N-1}{2N-1} p(x+1) = 2 \sum_{x=0}^N (N-x) p(x) \]
\[ \Rightarrow \left( \frac{2N+1}{2N+1} - \frac{1}{2N+1} \right) p(x+1) = 2 \sum_{x=0}^N (N-x) p(x) \]
\[ \Rightarrow E(X) - E(X) = 2N - 2E(X) \]
\[ \Rightarrow E(X) = -1. \]
Sum and Product Laws of Expectations:

- **Sum Law:** If \( X \) and \( Y \) are two discrete R.V.'s, then \( E(X+Y) = E(X) + E(Y) \).

  **Proof:** Let \( X \) take the values \( x_1, x_2, \ldots, x_i, \ldots \) and \( Y \) take the values \( y_1, y_2, \ldots, y_j, \ldots \), respectively.

  Define, \( \{ A_i \} \) and \( \{ B_j \} \) are two partitions of \( \Omega \).

  Now, \( E(X+Y) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (x_i + y_j) P[A_i \cap B_j] \)

  \[ = \sum_{i=1}^{\infty} x_i \sum_{j=1}^{\infty} y_j P[A_i \cap B_j] \]

  \[ = \sum_{i=1}^{\infty} x_i \sum_{j=1}^{\infty} y_j P[A_i \cap B_j] + \sum_{j=1}^{\infty} y_j \sum_{i=1}^{\infty} x_i P[A_i \cap B_j] \]

  \[ = \sum_{i=1}^{\infty} x_i P[A_i] + \sum_{j=1}^{\infty} y_j P[B_j] \]

  \[ = E(X) + E(Y). \]

Independence of Two Random Variables:

Consider two discrete R.V.'s \( X \) and \( Y \). Let \( \{ x_1, x_2, \ldots, x_i, \ldots \} \) and \( \{ y_1, y_2, \ldots, y_j, \ldots \} \) be the sets of most points of \( X \) and \( Y \), respectively.

Then, define \( A_i = \{ \omega : X(\omega) = x_i \} \) and \( B_j = \{ \omega : Y(\omega) = y_j \} \).

Here, \( \{ A_i \} \) and \( \{ B_j \} \) are two partitions of \( \Omega \).

**Definition:** The discrete R.V.'s \( X \) and \( Y \) are said to be independent iff \( \{ A_i \} \) and \( \{ B_j \} \) are two independent partitions of \( \Omega \).

iff \( P[A_i \cap B_j] = P[A_i] \cdot P[B_j] \) \( \forall (i, j) \)

i.e., \( P[X = x_i, Y = y_j] = P[X = x_i] P[Y = y_j] \) \( \forall i, j \).

- **Product Law:** If \( X \) and \( Y \) are independent discrete R.V.'s, then \( E(XY) = E(X) E(Y) \).

  **Proof:** Let \( \{ x_1, x_2, \ldots, x_i, \ldots \} \) and \( \{ y_1, y_2, \ldots, y_j, \ldots \} \) be the sets of most points of \( X \) and \( Y \), respectively.

  Then define, \( A_i = \{ \omega : X(\omega) = x_i \} \) and \( B_j = \{ \omega : Y(\omega) = y_j \} \).

  Hence, \( \{ A_i \} \) and \( \{ B_j \} \) are two partitions of \( \Omega \).
Since $X$ and $Y$ are independent,
\[ P\left[A_i \cap B_j\right] = P[A_i] P[B_j] \ \forall \ (i, j). \]

Now, \[ E(XY) = \sum_{i} \sum_{j} x_i y_j P\left[\{x, y\} = (x_i, y_j)\right] \]
\[ = \sum_{i} \sum_{j} x_i y_j P[A_i \cap B_j] \]
\[ = \sum_{i} \sum_{j} x_i y_j P[A_i] P[B_j], \text{ due to independence}. \]
\[ = \left\{ \sum_{i} x_i P[A_i] \right\} \left\{ \sum_{j} y_j P[B_j] \right\} \]
\[ = \left\{ \sum_{i} x_i P[X = x_i] \right\} \left\{ \sum_{j} y_j P[Y = y_j]\right\} \]
\[ = E(X).E(Y). \]

Remark: If \[ E(XY) = E(X).E(Y), \text{ then} \]
\[ \sum_{i} \sum_{j} x_i y_j P\left[X = x_i, Y = y_j\right] = \left\{ \sum_{i} x_i P[X = x_i] \right\} \left\{ \sum_{j} y_j P[Y = y_j]\right\} \]
\[ = \sum_{i} \sum_{j} x_i y_j P[X = x_i] P[Y = y_j] \]

This don't necessarily imply that
\[ P\left[X = x_i, Y = y_j\right] = P[X = x_i] P[Y = y_j] \ \forall \ i, j. \]
\[ \Rightarrow \ X \text{ and } Y \text{ are independent.} \]

Ex. Consider an RV $X \in \{1, 2, 3\}$. \[ P[X = -1] = \frac{1}{3}, P[X = 0] = \frac{1}{3}, P[X = 1] = \frac{1}{3}. \]
Define $Y = X^2$. Then show that \[ E(XY) = E(X)E(Y) \text{ but } X \text{ and } Y \text{ are not independent.} \]

Solution: \[ P[Y = 0] = P[X^2 = 0] = P[X = 0] = \frac{1}{3}, \]
\[ P[Y = 1] = P[X^2 = 1] = P[X = \pm 1] = \frac{2}{3}. \]
\[ E(X) = (-1) \cdot \frac{1}{3} + (+1) \cdot \frac{1}{3} = 0. \]
\[ E(XY) = (-1) \cdot 0 + 0 + 1 \cdot \frac{1}{3} = 0. \]
\[ E(Y) = 0 = E(X)E(Y). \]

But, \[ P[X = -1, Y = 1] = P[X = -1] = \frac{1}{3}, \]
\[ \neq P[X = -1] P[Y = 1] = \frac{1}{3} \cdot \frac{2}{3} = \frac{2}{9}. \]
\[ \Rightarrow \ X \text{ and } Y \text{ are not independent.} \]
Quantiles: A number $q_p$ satisfying $P[X \leq q_p] \geq p$ and $P[X > q_p] > 1 - p$, $0 < p < 1$, is called a $p^{th}$ order quantile of R.V. $X$.

If $q_p$ is a $p^{th}$ order quantile of an R.V. $X$, then:

$F(q_p) = p$ and $1 - F(q_p - 0) > 1 - p$,

i.e., $p > F(q_p - 0)$ and $p \geq F(q_p - 0)$

i.e., $F(q_p - 0) \leq p \leq F(q_p)$.

If $X$ is continuous R.V., then $P[X = q_p] = 0$, i.e., $F(q_p - 0) = F(q_p)$ and $F(q_p) = p$ and $q_p$ is the solution of the equation $F(x) = p$.

Ex. 1. Let $X$ be an R.V. with PMF:

$P[X = -2] = P[X = 0] = \frac{1}{4}$, $P[X = 1] = \frac{1}{3}$, $P[X = 2] = \frac{1}{6}$.

(i) Find Median ($q_{1/2}$)? (ii) Find a quantile of order $p = 0.2$ of the R.V. $X$.

Solution:

(i) $P[X \leq 0] = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$

Also, note that for any $x \in (0, 1)$,

$P[X \leq x] = \frac{1}{2}$ and $P[X > x] = \frac{1}{2} = 1 - \frac{1}{2}$.

It follows that every $x$, $0 \leq x < 1$ is a median of $X$.

(ii) $P[X \leq -2] = \frac{1}{4} > 0.2$ and $P[X > -2] = 1 > 1 - p$.

Hence, $p = 0.2$ is a quantile of $X$ is $-2$.

Ex. 2. Consider an R.V. $X$ with PDF:

$f(x) = \begin{cases} 0.2e^{-\theta(x-a)}, & \text{if } x > a \\ 0, \text{ otherwise} \end{cases}$

Find $q_{1/2}(X)$ and $E[X - q_{1/2}(X)]$.

Solution: Since $X$ is a continuous RV, the median $q_{1/2}$ is a solution of $F_X(x) = \frac{1}{2}$.

$F_X(x) = \int_{-\infty}^{x} f(t)dt = \frac{1}{2} = \int_{a}^{x} 0.2e^{-\theta(t-a)}dt = 1 - e^{-\theta(x-a)}$

$\Rightarrow \theta(x-a) = -\ln 2$,

$\Rightarrow x = a + \frac{1}{\theta}\ln 2$. 
Hence, \( E_{1/2} = a + \frac{\alpha}{\beta} \ln 2 \).

Now, \( E \left| X - E_{1/2} \right| = \int_a^\infty \left| x - \alpha - \frac{\ln 2}{\beta} \right| \cdot e^{-\beta (x - \alpha)} \, dx \)

\[ \frac{\ln 2}{\beta} \cdot \int_a^\infty \frac{1}{y} \cdot e^{-\theta (y + \frac{\ln 2}{\beta})} \, dy \]  

\[ = \frac{\ln 2}{\beta} \cdot \int_a^\infty \frac{1}{y} \cdot e^{-\theta y} \, dy \]  

\[ = \frac{e^{-\ln 2/\beta}}{1/2} \]  

\[ = \frac{1}{2} \int_0^{\infty} z e^{2} \, dz \]  

\[ = \frac{1}{2} \int_0^{\infty} 2 e^{-2z} \, dz + \int_{\ln 2}^{\infty} e^{-2z} \, dz \]  

\[ = \frac{1}{2} \left\{ \left[ \left( e^{-2} \right)^{0} \right]_{-\infty}^{0} + \Gamma(2) \right\} \]  

\[ = \frac{1}{2} \left\{ \left[ -1 + 2 \ln 2 - 2 + 1 \right] \right\} \]  

\[ = \frac{\ln 2}{\beta} \]  

\[ = \left( E_{1/2} - \alpha \right) \]

---

**Measures of Central Tendency:**
- \( \mu = E(X) \), \( E_{1/2}(X) \) are the measures of central tendency of the distribution of \( X \).
- **Mode:** If \( X \) is a discrete (continuous) RV, then the value \( \alpha \) for which the PMF (or PDF) \( f_X(x) \) is maximum, is called the mode of the distribution of the RV \( X \).
- **Harmonic Mean:** HM of a non-zero RV is given by \( \text{HM} = \frac{1}{E\left(\frac{1}{X}\right)} \), provided the expectation exists.
- **Geometric Mean:** GM of a positive RV \( X \) is denoted by \( G \) and it is given by \( \log G = E(\log X) \).

**Measure of Dispersion:**
- \( SD(X) = \sigma_X = \sqrt{V(X)} \); \( CV = \frac{\sqrt{V(X)}}{E(X)} \), provided \( E(X) > 0 \).

**Measure of Skewness and Kurtosis:**
- \( \gamma_1 = \frac{\mu^3}{\mu^2} \) is a measure of skewness.
- \( \beta_1 = \frac{\mu_3^2}{\mu_2^3} \).
- \( \gamma_2 = \beta_2 - 3 = \frac{\mu_4}{\mu_2^2} - 3 \) is a measure of excess of kurtosis of \( X \).
Ex. Let \( F(t) \) be the probability that a system fails by time \( t \) and let \( \gamma(t) \) be the probability of failure in the interval \((t, t+\Delta t)\) given that it has survived up to \( t \). S.T.\, F(t) satisfies the differential equation \( F'(t) = \gamma(t) \), leading to the functional form \( F(t) = 1 - \exp\left[ \int_0^t \gamma(x) \, dx \right] \). In particular, if \( \gamma(t) = 0 \), find \( F(t) \).

Solution: Let \( T \) be the lifetime of the system. Hence, \( F(t) = P[T \leq t] \).

By problem, \( P\left[ t < T < t+\Delta t \mid T > t \right] = \gamma(t) \Delta t + O(\Delta t) \)

\[ \Rightarrow \frac{P\left[ t < T < t+\Delta t \right]}{P[T > t]} = \gamma(t) \Delta t + O(\Delta t) \]

\[ \Rightarrow \frac{F(t+\Delta t) - F(t)}{\Delta t} = \gamma(t) \Delta t + O(\Delta t) \]

\[ \Rightarrow \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t} = \gamma(t) + \lim_{\Delta t \to 0} \frac{O(\Delta t)}{\Delta t} \]

\[ \Rightarrow \frac{F'(x)}{1-F(x)} = \gamma(x), \text{ for } x > 0. \]

\[ \Rightarrow \int_0^t \frac{F'(x)}{1-F(x)} = \int_0^t \gamma(x) \, dx \]

\[ \Rightarrow \left[ -\log_e (1-F(x)) \right]_0^t = \int_0^t \gamma(x) \, dx \]

\[ \Rightarrow -\log_e \left( \frac{1-F(t)}{1-F(0)} \right) + \log_e \left( \frac{1-F(0)}{1-F(t)} \right) = \int_0^t \gamma(x) \, dx \]

\[ \Rightarrow \log_e \frac{1-F(t)}{1-F(0)} = \int_0^t \gamma(x) \, dx \quad \text{As, } T \text{ is a non-negative RV, } F(0) = 0 \]

\[ \Rightarrow F(t) = 1 - e^{-\int_0^t \gamma(x) \, dx} = 1 - e^{-\theta t}, \quad t > 0. \]

Remark:

\[ \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t} = \frac{F'(t)}{1-F(t)} \]

is called the instantaneous failure rate or hazard rate at time point \( t \).
Symmetric Distribution:

Definition: An R.V. $X$ is said to be symmetrically distributed about $a$ if
\[ P[X \leq a-x] = P[X > a+x] \quad \forall x. \]
\[ F_X(a-x) = 1 - F_X(a+x) + P[X = a+x] \quad \forall x. \]

For a discrete R.V. $X$ with PMF $f(x)$, $X$ is said to be symmetric about $a$ if
\[ f(a-x) = f(a+x) \quad \forall x. \]

For a continuous R.V. $X$ with PDF $f(x)$, we have
\[ F_X(a-x) = 1 - F_X(a+x) \quad \text{as} \quad P[X = a+x] = 0. \]
\[ f(a-x) = f(a+x) \quad \forall x. \]

Result: Show that for a symmetric R.V., all odd order central moments about zero, provided they exist. Also show that the mean of the R.V. is the point of symmetry.

Proof: Let $X$ be a continuous R.V. with DF $F(x)$ and let $X$ be symmetric about a point $a$. By definition, $f(a-x) = f(a+x) \quad \forall x$. Now,
\[
E(X-a)^{2n-1} = \int (x-a)^{2n-1} f(x) \, dx
\]
\[
= \int_a^\infty (x-a)^{2n-1} f(x) \, dx + \int_\infty^a (x-a)^{2n-1} f(x) \, dx
\]
\[
= \int_0^\infty (a+y)^{2n-1} f(a+y) \, dy + \int_\infty^a y^{2n-1} f(a+y) \, dy \quad \text{[:: y = x-a]} \]
\[
= \int_0^\infty (-u)^{2n-1} f(a+u) \, du + \int_\infty^a u^{2n-1} f(a+u) \, du \quad \text{[:: f(a-u) = f(a+u)]}
\]
\[
= 0, \quad \forall n \in \mathbb{N}. \]

For $n = 1$, $E(X-a) = 0 \Rightarrow E(X) = a$.

Hence, $\mu_{2n-1} = E(X-X(E(X)))^{2n-1} = E(X-a)^{2n-1} = 0 \quad \forall n \in \mathbb{N}$.

Result: If $X$ is symmetrically distributed about $a$, show that median of $X$ is $a$.

Proof: By definition of symmetric distribution about $a$,
\[ P[X \leq a-x] = P[X > a+x] \quad \forall x. \]

For $x = 0$, we have $P[X \leq a] = P[X > a]$. 

Scanned by CamScanner
Ctanjul Classes
Page No. 103
Problem: (Banach's Match Box Problem) A certain mathematician carries two match boxes in his pocket. Each time he wants to use a match, he selects one of the boxes at random. Find the probability of selecting a match from either box. He discovers that one box is empty, then the other box contains $n$ matches, $n = 0, 1, 2, \ldots, n$. Suppose $n$ is the no. of matches in the other box.

Solution: Since there is one match-box in his two pockets, the probability of selecting at random to have a match is $p = \frac{1}{2}$. Let us identify 'success' with the choice of the left pocket. Then the left pocket will be found empty at a moment when the right pocket contains exactly $n$ matches, if exactly $(N-n)$ failures preceding the $(N+1)$th success. The probability of this event is

$$P_L = \binom{2n-n}{n-n} \left(\frac{1}{2}\right)^n \left(\frac{1}{2}\right)^{n-n} \cdot \frac{1}{2} = \binom{2n-n}{n-n} \cdot \frac{1}{2^{2n-n+1}}.$$

The required event is that $(2n-n+1)$ trials is needed to get the $(n+1)$th success. For the occurrence of this event we must have a success at the last trial and there are $(N-n)$ failures in the first $2n-n)$ trials.

The same argument applied to the right pocket and then the corresponding probability is

$$P_R = \binom{2n-n}{n-n} \cdot \frac{1}{2^{2n-n+1}}.$$

Hence, the probability that there are $n$ matches in one box when the other box is found empty is $P_L + P_R = 2 \binom{2n-n}{n-n} \cdot \frac{1}{2^{2n-n+1}} = \binom{2n-n}{n-n} \cdot \frac{1}{2^{2n-n}}$.

Note that, if $R$ be the expected no. of matches, then we have

$$n - E(R) = n - \sum_{n=0}^{\infty} n \cdot P[R = n]$$

$$= \sum_{n=0}^{\infty} \binom{2n-n}{n-n} \cdot \frac{1}{2^{2n-n}}$$

$$= \sum_{n=0}^{\infty} \left(\frac{(2n+1)-(n+1)}{2}\right) \binom{2n-n-1}{n-n-1} \cdot \frac{1}{2^{2n-n-1}}$$

$$= \frac{2n+1}{2} \sum_{n=0}^{\infty} P[R = n+1] - \frac{1}{2} \sum_{n=0}^{\infty} \frac{(n+1)P[R = n+1]}{2}$$

$$= \frac{2n+1}{2} \left[ 1 - P[R = 0] \right] - \frac{1}{2} E(X) = \frac{2n+1}{2} \left[ 1 - \binom{2n}{n} \frac{1}{2^{2n-1}} \right] - E(0)$$

$$\Rightarrow \frac{1}{2} E(R) = \frac{1}{2} \left(\binom{2n+1}{n} \frac{1}{2^{2n-1}} \right) \left[ 1 - \binom{n}{n} \frac{1}{2^{2n-1}} \right] \Rightarrow E(R) = \binom{2n+1}{n} \frac{1}{2^{2n-1}} - 1.$$
**JENSEN'S INEQUALITY:**

Convex Function:- (i) If $f(x)$ is twice differentiable, i.e. $f''(x)$ exists and $f''(x) \geq 0 \ \forall \ x \in I$, then $f(x)$ is convex in the interval $I$.

(ii) A function $f(x)$ is said to be convex on an interval $I$ if for $x_1, x_2 \in I$,

\[ f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) ; 0 \leq \lambda \leq 1. \]

Jensen's Inequality:- If $f(x)$ is continuous and convex function on $I$ and $X$ is an RV, such that $P[X \in I] = 1$, then $\mathbb{E}[f(X)] \geq f(\mathbb{E}(X))$.

Proof:- Assuming $f(x)$ is twice differentiable. By Taylor's theorem,

\[ f(x) = f(\mu) + (x-\mu)f'(\mu) + \frac{(x-\mu)^2}{2!} f''(\mu^*) \]

where $\mu^*$ lies between $\mu$ and $x$; $\mu = \mathbb{E}(X)$.

Note that, as $f(x)$ is convex, then $f''(x) \geq 0 \ \forall \ x$. Thus

\[ f''(\mu^*) \geq 0 \]

\[ \Rightarrow f(x) \geq f(\mu) + (x-\mu)f'(\mu) ; \forall x > 0 \]

For an RV $X$,

\[ f(X) \geq f(\mu) + (X-\mu)f'(\mu) \]

\[ \Rightarrow \mathbb{E}(f(X)) \geq f(\mu) + \mathbb{E}(X-\mu)f'(\mu) \]

\[ \Rightarrow \mathbb{E}(f(X)) \geq f(\mathbb{E}(X)). \]

Remark:- A function $f(x)$ is concave on $I$ if $-f(x)$ is convex on $I$.

Jensen's Inequality:- For a concave function $f(x)$ on $I$, $E(f(X)) \leq f(E(X))$.

Proof:-

$f(x)$ is concave,

\[ \Rightarrow -f(x) \text{ is convex}. \]

\[ \Rightarrow E(-f(X)) \geq -f(E(X)). \]

\[ \Rightarrow E(f(X)) \leq f(E(X)). \]
Ex. For an RV $X$ which assumes only positive values, show that:

(i) $E\left(\frac{1}{X}\right) > \frac{1}{E(X)}$

(ii) $E(\log_e X) \leq \log_e E(X)$

Solution:

(i) Let $f(x) = \frac{1}{x}, x > 0$

By Taylor's theorem, $f(x) = f(\mu) + (x - \mu) f'(\mu) + \frac{(x - \mu)^2}{2!} f''(\mu)$.

$\mu = E(X)$, $\mu$ lies between $\mu$ and $x$.

$f'(x) = -\frac{1}{x^2}$, $f''(x) = \frac{2}{x^3} > 0 \forall x > 0$.

Hence, $f(x) > f(\mu) + (x - \mu) f'(\mu)$

$\Rightarrow E(f(x)) > E(\mu)$

$\therefore E\left(\frac{1}{X}\right) > \frac{1}{E(X)}$

(ii) Take $f(x) = \log_e x$, $x > 0$, $E(X) = \mu$.

$f'(x) = \frac{1}{x}$, $f''(x) = -\frac{1}{x^2} < 0$.

Jensen's inequality states that $E(f(x)) \leq f(E(X)) = f(\mu)$

$\Rightarrow E(\log_e X) \leq \log_e E(X)$

2. CAUCHY-SCHWARTZ INEQUALITY:

$E\left(g^2(X)\right) E\left(h^2(X)\right) \geq E\left(g(X) h(X)\right)$, provided $E(g(X))$ and $E(h(X))$ both exist.

Proof: For any real $\lambda$, $E\left(g(X) + \lambda h(X)\right)^2 \geq 0$

$\Rightarrow E\left(g^2(X)\right) + \lambda^2 E\left(h^2(X)\right) + 2\lambda E\left(g(X) h(X)\right) \geq 0$

$\Rightarrow E\left(g^2(X)\right) + 2\lambda E\left(g(X) h(X)\right) + \lambda^2 E\left(h^2(X)\right) \geq 0$

$\Rightarrow \lambda^2 + b\lambda + c \geq 0 \quad \left[ \text{Take } a = E\left(h^2(X)\right), b = 2 E\left(g(X) h(X)\right), c = E\left(g^2(X)\right) \right]$

Choose $\lambda = -\frac{b}{2a}$.

$\Rightarrow E\left(h^2(X)\right) E\left(g^2(X)\right) \geq E^2\left(h(X) g(X)\right)$

Note: In general, for two jointly distributed RVs $X$ and $Y$,

$E\left(g^2(X)\right) E\left(h^2(Y)\right) \geq E\left(g(X) h(Y)\right)$, provided $E\left(g^2(X)\right)$ and $E\left(h^2(Y)\right)$ both exist.

Holds in Cauchy-Schwarz Inequality, i.e., $g(X) h(X) = 0$ almost everywhere.
Application of C.S. Inequality:

(a) \[ E(g^2(X))E(h^2(Y)) \geq E(g(X)h(Y)) \]
Choose \( g(X) = X \), \( h(X) = 1 \), almost everywhere.
\[
E(X^2) \geq E^2(X) \quad \Rightarrow \quad E(X^2) - E^2(X) \geq 0 \quad \Rightarrow \quad \text{Var}(X) \geq 0.
\]

(b) Replace 'X' by \( |X - E(X)| \) in (a),
\[
E(\sqrt{E(X)} - E(X))^2 \geq E(\sqrt{E(X)} - E(X))^2.
\]
\[
\Rightarrow \quad \sqrt{E(X)} - E(X) \geq E(\sqrt{X} - E(X)) \]
\[
\Rightarrow \quad \text{SD}(X) \geq \text{MD}(X).
\]

(c) Replace 'X' by \( \sqrt{X} \) in (a), where \( P(X > 0) = 1 \),
\[
E(X) \geq E(X^2) \quad \Rightarrow \quad \sqrt{E(X)} \geq E(\sqrt{X}) \]
(d) Let \( g(X) = \sqrt{X} \), \( h(X) = \frac{1}{\sqrt{X}} \), \( P(X > 0) = 1 \).
\[
E(X) E\left(\frac{1}{X}\right) \geq 1 \]
\[
\Rightarrow \quad E(X) \geq \frac{1}{E\left(\frac{1}{X}\right)} \quad \Rightarrow \quad \text{AM} \geq \text{HM}.
\]

(e) \( g(X) = X - E(X) \), \( h(Y) = Y - E(Y) \).
Then \( E((X - E(X))^2) \geq E((Y - E(Y))^2) \).
\[
\Rightarrow \quad \text{Var}(X) \geq \text{Var}(Y) \]
\[
\Rightarrow \quad |\text{Cov}(X,Y)| \leq 1
\]

Note: Let \( X \) be an RV with mean \( \mu \) and variance \( \sigma^2 > 0 \).
Define, \( Z = \frac{X - \mu}{\sigma} \). Note that, \( E(Z^2) = \frac{E((X-\mu)^2)}{\sigma^2} \)
\[
= \frac{\mu^2}{\mu^2} = \frac{\sigma}{\sqrt{\mu^2}}.
\]
Problem: Show the following:
(i) \( \beta_2 > 1 \)
(ii) \( \beta_2 > \beta_1 \)
(iii) \( \beta_2 > \beta_1 + 1 \)

Solution:

(i) Take, \( q(z) = z^2, \ h(z) = z \).

\( \Rightarrow E(z^4) > E^2(z^2) \), provided \( E(z^4) \) exists, i.e., \( \frac{\mu_4}{\mu_2^2} \) exists.

Now, \( E(z^4) = \frac{\mu_4}{\mu_2^2}, \ E(z^2) = \frac{\mu_2}{\mu_2} = 1 \).

\( \Rightarrow \ \frac{\mu_4}{\mu_2^2} > 1 \Rightarrow \ \beta_2 > 1 \).

\( \Rightarrow \) holds iff \( z^2 = 0 \) almost everywhere.

\( \Rightarrow z = k \) a.e.

i.e. \( \left| \frac{x - \mu}{\sigma} \right| = k \), \( \Rightarrow x = \mu \pm k\sigma \).

i.e. \( x \) assumes too distinct values with equal probability.

(ii) \( q(z) = z^2, \ h(z) = z \).

\( \Rightarrow E(z^4) E(z) > E^2(z^3) \)

\( \Rightarrow \frac{\mu_4}{\mu_2^2} > \frac{\mu_3^2}{\mu_2^3} \).

i.e. \( \beta_2 > \beta_1 \).

\( \Rightarrow \) holds iff \( q(z) = k^2 h(z) \) a.e.

i.e. \( z \) is degenerate r.v.

i.e. \( x \) is also degenerate r.v.

in that case, \( \mu_2 \) vanishes.

Thus, equality will not hold good, i.e., \( \beta_2 > \beta_1 \).

(iii) \( q(z) = z^2 - 1, \ h(z) = z \).

\( \Rightarrow E(z^4 - 2z^2 + 1) E(z^2) > E(z^3 - z) \).

\( \Rightarrow \frac{\mu_4}{\mu_2^2} - 1 > \frac{\mu_3^2}{\mu_2^3} \).

\( \Rightarrow \ \beta_2 > \beta_1 + 1 \).

\( \Rightarrow \) holds iff \( q(z) = k h(z) \) a.e.

\( \Rightarrow z^2 - 1 = k \).

\( \Rightarrow z \) assumes too distinct values.

\( \Rightarrow x \) assumes too distinct values not necessarily symmetrically placed const. \( \mu \).
Moment Generating Function:

Given the form $\Psi(t,x) = e^{tx}$, by the m.g.f. of a r.v. $x$, we mean $E(e^{tx})$, provided this expectation exists for all $t$ satisfying $|t| < h$, i.e., $-h < t < h$, $h > 0$.

It is denoted by $M(t)$ or $M_x(t)$.

Example:

(i) Let $x \sim \text{Bin}(n,p)$, then $E(e^{tx})$ exists for all.

Hence, the m.g.f. exists and is given by

$$M_x(t) = E(e^{tx})$$
$$= \sum_{x=0}^{n} e^{tx} \binom{n}{x} p^x q^{n-x}$$
$$= \sum_{x=0}^{n} \left( \binom{n}{x} p e^t \right) q^{n-x}$$
$$= \left( q + pe^t \right)^n = (q + pe^t)^n$$

(ii) Let $x \sim \text{Poiss}(\lambda)$, then $E(e^{tx})$ exists for all $t$.

Hence, the m.g.f. is defined and given by

$$M_x(t) = E(e^{tx})$$
$$= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda} \lambda^x}{x!}$$
$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$
$$= e^{-\lambda} e^{\lambda e^t}$$
$$= e^{-\lambda} (1 - e^t) = e^{\lambda (e^t - 1)}$$
USES:

1. Suppose all moments of \( x \) exist, then expanding \( M_x(t) \) in a power series of \( x \), we get:

\[
M_x(t) = E(e^{tx})
\]

\[
= \sum_{x=0}^{\infty} e^{tx} \cdot P(x = x)
\]

\[
= \sum_{x=0}^{\infty} \left( \sum_{k=0}^{\infty} \frac{(tx)^k}{k!} \right) P(x = x)
\]

\[
= \sum_{x=0}^{\infty} \sum_{k=0}^{\infty} \frac{t^k x^k P(x = x)}{k!}
\]

\[
= \sum_{k=0}^{\infty} \frac{t^k}{k!} \sum_{x=0}^{\infty} x^k P(x = x)
\]

Since the inner sum is finite.

Thus the \( r \)th raw moment of \( x \) is obtainable as the coefficient of \( \frac{t^r}{r!} \) in the power series expansion of \( M_x(t) \).

Hence, the name 'm.g.f.'

The same thing can be shown for continuous case also.

Central moment generating function:

\[
M_{x - \mu}(t) = E(e^{t(x - \mu)}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot E(x^k).
\]

Now,

\[
M_{x - \mu}(t) = E(e^{t(x - \mu)}) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \cdot E(x^k).
\]

The m.g.f. of a linear \( Y = a + bx \) is

\[
M_{a+bx}(t) = \int_{-\infty}^{\infty} e^{t(a+bx)} dF(x)
\]

\[
= e^{at} M_x(bt),
\]

which is obtained by replacing \( t \) by \( bt \) in \( M_x(t) \) and multiplying the result by \( e^{at} \).
2. Let $X$ and $Y$ are independent random variables, with m.g.f. $M_X(t)$ and $M_Y(t)$ for $t$ satisfying

- $-h_0 < t < h_0$ where $h_0 = \min(h_1, h_2)$ and $M_X(t)$ is defined for $-h_1 < t < h_1$.
- $M_Y(t)$ is defined for $-h_2 < t < h_2$.

Then, the m.g.f. of $X + Y$ is given by

$$M_{X+Y}(t) = E(e^{t(X+Y)}) = E(e^{tX}e^{tY})$$

$$= E(e^{tX}) \cdot E(e^{tY}) \quad [\text{since } X \text{ and } Y \text{ are independent}].$$

$$= M_X(t) \cdot M_Y(t).$$

Defined for $-h_0 < t < h_0$.

Hence knowing the m.g.f. of $X$ and $Y$, one can obtain the m.g.f. of $(X+Y)$ provided $X$ and $Y$ are independent.

Since the m.g.f. is unique, one can then obtain the distribution of $(X+Y)$.

Example: Let $X \sim \text{Bin}(m_1, p)$ and $Y \sim \text{Bin}(m_2, p)$ independent.

- $M_X(t) = (1 - pt)^{m_1}$
- $M_Y(t) = (1 - pt)^{m_2}$

Then,

$$M_{X+Y}(t) = M_X(t) \cdot M_Y(t)$$

$$= (1 - pt)^{m_1+m_2},$$

which is the m.g.f. of a $\text{Bin}(m_1+m_2, p)$ variable.

Hence, $X + Y \sim \text{Bin}(m_1+m_2, p)$.

Find the m.g.f.:

(i) $f(x) = \frac{e^{-\theta} \theta^x}{x!}$, $0 < \theta < \infty$

(ii) $f(x) = \frac{1}{B(m, m)} x^{m-1} (1-x)^{m-1}$, $0 < x < 1$.  

(iii) $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $-\infty < x < \infty$
3. The difficulty with mgf is that it may not always exists. Suppose, \( x \) is a non-negative r.v. for which for some positive integer \( n \) \( E(x^n) \) does not exist. Show that in this case, mgf does not exist.

\[
e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!}
\]

for \( t > 0 \), \( e^{tx} > t^n \cdot \frac{a^n}{n!} \)

\[
\Rightarrow \int_0^{\infty} e^{tx} f(x) \, dx > \int_0^{\infty} t^n \cdot \frac{a^n}{n!} f(x) \, dx
\]

\[
= \frac{t^n}{n!} \int_0^{\infty} a^n f(x) \, dx
\]

Since, \( E(x^n) \) does not exist, \( \int_0^{\infty} a^n f(x) \, dx = \infty \).

\[
\Rightarrow E(e^{tx}) \text{ does not exist for } t > 0.
\]

Hence the mgf does not exist.

- Ex. i) Cauchy distribution, \( f(x) = \frac{1}{\pi} \cdot \frac{1}{1 + x^2} \), \(-\infty < x < \infty\),
  \( \Rightarrow E(x) \) doesn't exist. \( \Rightarrow \) mgf does not exist.

ii) Lognormal distribution, mgf does not exist.

**MGf in bivariate case:**

Let \( x = (x_1, x_2) \) is a bivariate random variable.

Then the mgf of \( x \) is defined as

\[
M_{x_1, x_2}(t_1, t_2) = E(e^{t_1 x_1 + t_2 x_2})
\]

where \( t = (t_1, t_2) \)

Now, the mgf of \( x_1 \) is obtained by putting \( t_2 = 0 \)

\[
M_{x_1, x_2}(t_1, 0) = E(e^{t_1 x_1})
\]

\[
= M_{x_1}(t_1).
\]

Note that, here, \( M(t_1, t_2) = E(e^{t_1 x_1 + t_2 x_2}) \) for all values \( t_1 \) and \( t_2 \), is the moment generating function of the joint distribution of \( x_1 \) & \( x_2 \),

\[
E(X_i | x_j) = \frac{\partial^2 M(0, 0)}{\partial t_1 \partial t_2}.
\]
Theorem: The m.g.f. of the sum \( S_n = X_1 + \ldots + X_n \) of \( n \) independent r.v.'s \( X_1, X_2, \ldots, X_n \) is 
\[
M_{S_n}(t) = M_{X_1}(t) \cdots M_{X_n}(t),
\]
where \( M_{X_i}(t) \) is the p.m.f. of \( X_i \) provided all the m.g.f.'s exist.

Proof: 
\[
M_{S_n}(t) = E(e^{tS_n}) = E\left(\prod_{i=1}^{n} e^{tX_i}\right) = \prod_{i=1}^{n} E(e^{tX_i}),
\]
as \( X_i \)'s are independent.

Example: \( X \) is normally distributed with mean \( \mu \) and s.d. \( \sigma \).

\[
f(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad -\infty < x < \infty
\]

Then m.g.f. of \( X \) about mean(\( \mu \)) is 
\[
M_{X-\mu}(t) = E(e^{t(X-\mu)})
\]
\[
= \frac{1}{\sigma \sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{t(x-\mu)^2}{2\sigma^2}} dx
\]
\[
= e^{\frac{t\mu^2}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2\sigma^2}[(x-\mu)^2+\frac{t\sigma^2}{2}]} dx
\]
\[
= e^{\frac{t\mu^2}{2}} \sum_{n=0}^{\infty} \left(\frac{t\sigma^2}{2}\right)^n / n!
\]
Thus, 
\[
E(X-\mu) = 0 \quad \text{on}, \quad E(X) = \mu.
\]
And 
\[
E((X-\mu)^{2n+1}) = 0, \quad E((X-\mu)^{2n}) = \mu^{2n} = \frac{(2n)!}{2^n x^{2n}}
\]
\[
e (2n-1)(2n-3) \cdots \times 1 \times 0^{2n}
\]
for \( n = 1, 2, \ldots \)

Note: The relation between the p.g.f. \( P_x(t) \) the m.g.f. of a non-negative integer-valued random variable \( X \) is given by 
\[
M_X(t) = P_X(e^t).
\]
If a distribution is specified by its m.g.f., then the moments (about origin) can be obtained as follows:
\[
\mu_k' = \left. \frac{d^k M_X(t)}{dt^k} \right|_{t=0}
\]
8.29. If \( X \) has the negative binomial distribution with

\[
f(x) = \binom{x+n-1}{x} p^n q^x \quad \text{if} \quad x \geq 0
\]

\[= 0 \quad \text{otherwise}
\]

where \( n \) is a positive integer, \( 0 < p < 1 \), and \( q = 1-p \), S.T.,
the m.g.f. of \( X \) exists. Also obtain the m.g.f.

**Ans:**

\[
M_x(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} f(x)
\]

\[
= p^n \sum_{x=0}^{\infty} \binom{x+n-1}{x} q^x e^{tx}
\]

is an absolutely convergent series for \( qe^t < 1 \) i.e. for \( t > -\ln q \).

Hence the m.g.f. exists. Also,

\[
M_x(t) = p^n (1 - qe^t)^{-n}, \quad \text{for} \quad t < -\ln q.
\]

**Alternative:**

we know,

\[
M_x(t) = P_X(e^t)
\]

we know,

\[
P_x(t) = p^n (1 - at)^{-n}
\]

for \( |at| < 1 \)

Now,

\[
M_x(t) = p^n (1 - qe^t)^{-n},
\]

for \( t < -\ln q \).

8.30. Suppose \( X \) is a non-negative r.v. for which some positive integer \( n \), \( E(X^n) \) does not exist. S.T. \( E(e^{tx}) \) exists if \( t \leq 0 \). Does the m.g.f. exist?

**Ans:**

\[
e^{tx} = \sum_{n=0}^{\infty} \frac{(tx)^n}{n!}
\]

for \( t > 0 \), then,

\[
\exp(tx) > \frac{t^n}{n!} \cdot 2^{tn} \quad (\text{for} \quad x > 0)
\]

\[
\Rightarrow \int \exp(tx) dF(x) = \infty, \quad \text{since} \quad \int x^n dF(x) = \infty
\]

On the other hand, for \( t \leq 0 \),

\[
\Rightarrow 0 \leq \exp(tx) \leq 1
\]

\[
\Rightarrow 0 \leq \int \exp(tx) dF(x) \leq 1
\]

As such, \( E(e^{tx}) \) is defined for \( t \leq 0 \). Hence, the p.m.f. does not exist, since we get any

\[
h > 0 \Rightarrow \int_{-\infty}^{\infty} e^{tx} f(x) dx < \infty \quad \text{for} \quad |t| \leq h
\]
8.31. S.t. the m.g.f. does not exist for the continuous distribution with p.d.f.

\[ f(x) = \begin{cases} \frac{1}{4x^5} & \text{if } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases} \]

**Ans:** In this case,

\[ \int \limits_0^\infty |x^5| f(x) \, dx = \frac{1}{4} \int \limits_0^\infty x \, dx = \infty \]

Hence, \( E(x^5) \) does not exist.

\[ \Rightarrow E(e^{tx}) \text{ does not exist for } t > 0. \]

Hence, m.g.f. does not exist.

8.32. S.t. the m.g.f. does not exist for a Cauchy distribution.

**Ans:**

\[ f(x) = \frac{1}{\pi} \frac{1}{1 + x^2}, \quad -\infty < x < \infty \]

\[ E|x|^\rho = \frac{2}{\pi^2} \int \limits_0^\infty x^\rho \frac{1}{1 + x^2} \, dx \]

\[ = \frac{1}{\pi} \int \limits_0^\infty 2^{1 - \frac{\rho}{2}} (1 - e^{-\frac{\pi}{2} x^2}) \right[ \frac{(\rho + 1)/2}{\pi} \left] - 1 \, dx \]

which is a beta function \( \Phi(\rho) \) exists for \( \rho < 1 \)

and diverges for \( \rho > 1 \).

\[ E|x|^\rho \text{ does not exist for } \rho > 1 \]

\[ \Rightarrow E(e^{tx}) \text{ does not exist for } \rho > 1 \]

\[ \Rightarrow \text{m.g.f. does not exist.} \]

**Alt. method:**

Now, if \( t > 0 \),

\[ \int \limits_{-\infty}^{\infty} e^{tx} f(x) \, dx > \int \limits_{0}^{\infty} e^{tx} f(x) \, dx > \left[ \frac{t}{1 + x^2} \right]_0^{\infty} = \frac{\pi}{2} \left[ \ln(1 + x^2) \right]_0^{\infty} = \frac{t}{2\pi} \ln(1 + x^2) \]

For \( t > 0 \), m.g.f. does not exist.

If \( t < 0 \),

\[ 0 \leq \int \limits_{-\infty}^{\infty} e^{tx} f(x) \, dx \leq 1, \quad = \infty. \]
\[ \frac{d}{dt} x, \frac{d}{dx} x \]
8.8.4. If $X$ has the Laplace distribution with p.d.f. 
\[ f(x) = \frac{\alpha}{2} e^{-\alpha |x|}, \quad -\infty < x < \infty, \quad \alpha > 0, \]
find the m.g.f. of $X$. For what value of $t$ is it defined.

Hence obtain $E(X)$ and $\text{Var}(X)$.

**Answer**

\[
E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx
\]

\[
= \frac{\alpha}{2} \int_{-\infty}^{\infty} e^{tx} e^{-\alpha |x|} dx
\]

\[
= \frac{\alpha}{2} \left[ \int_{-\infty}^{0} e^{tx} e^{-\alpha x} dx + \int_{0}^{\infty} e^{tx} e^{-\alpha x} dx \right]
\]

The integrals are convergent for $|t| < \alpha$; defined for $|t| < \alpha$.

As such $E(e^{tx})$ exists for $|t| < \alpha$, implying that the m.g.f. exists.

\[
M_X(t) = E(e^{tx}) = \frac{\alpha}{2} \left[ \frac{1}{t+\alpha} + \frac{1}{\alpha-t} \right]
\]

\[
= \frac{1}{2} \left[ (1+\frac{t}{\alpha})^{-1} + (1-\frac{t}{\alpha})^{-1} \right]
\]

\[
= \frac{1}{2} \left[ (1+\frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \cdots) + (1+\frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \cdots) \right]
\]

\[
= \left[ 1 + \frac{t}{\alpha} + \frac{t^2}{\alpha^2} + \frac{t^3}{\alpha^3} + \cdots \right]
\]

\[
M_X(t) = \frac{1}{1 - \frac{t}{\alpha}} = \frac{\alpha}{\alpha - t}
\]

\[
M_X'(t) = \frac{2t \alpha}{(\alpha - t)^2}
\]

\[
M_X''(t) = \frac{2 \alpha^2 (\alpha - t) - 2 \alpha^2 t}{(\alpha - t)^4}
\]

\[
E(X) = M_X'(0) = 0
\]

\[
E(X^2) = M_X''(0) = \frac{2 \alpha^2}{\alpha^2} = \frac{2}{\alpha^2}
\]

\[
\text{Var}(X) = E(X^2) - [E(X)]^2 = \frac{2}{\alpha^2}
\]
Cumulant Generating function:

Suppose for a r.v. $X$, the m.g.f. $M_X(t)$ is defined then $M_X(t)$ is also defined and is called the c.g.f. of $X$.

It is denoted by $K_X(t)$.

If the c.g.f. is expanded as a power series in $t$ then the co-efficient of $\frac{t^n}{n!}$ in that series is called the $n$th cumulant of $X$, it is denoted by $\kappa_n$.

Characteristic function: — We have already noted that the m.g.f. of a r.v. does not always exist, in such cases also the c.f. exists.

Suppose $Z = X + iY, \quad i = \sqrt{-1}$

Then we say that, $E(Z) = E(X) + iE(Y)$

If $X$ is an ordinary r.v., then $E(e^{itx})$ is called the characteristic function of $X$ and is denoted by $\phi_X(t)$.

$\int_{-\infty}^{\infty} |e^{itx}|dF(x) < \infty$, since

$|e^{itx}| = |\cos x + is \sin x| = \sqrt{\sin^2 x + \cos^2 x} = 1,$

$\int_{-\infty}^{\infty} |e^{itx}|dF(x) = \int_{-\infty}^{\infty} dF(x) = 1, \quad i.e., \text{finite.}$

$\phi_X(t) = E(e^{itx})$ always exist.

Alternative Definition of CF:

Let $X$ be an r.v. The complex valued function $\Phi$ defined on $\mathbb{R}$ by

$\Phi(t) = E(e^{itx}) = E(e^{itX}) + iE(sintx), \quad t \in \mathbb{R}$

where $i = \sqrt{-1}$ is the imaginary unit, is called the characteristic function (CF) of r.v. $X.$
Some Probability Inequalities

The inequalities which contain probability in either left side or right side or in both side, are called "Probability Inequalities".

Markov's Inequality:

Statement: Let $X$ be a r.v. having finite expectation, i.e., $E(X)$ converges. Then for any non-negative quantity $a$, we have the inequality:

$$P(X > a) \leq \frac{E(X)}{a}.$$  

Proof: Let us define a r.v. $Y$ such that

$$Y = \begin{cases} a & \text{if } X > a \\ 0 & \text{otherwise} \end{cases}$$

$$X > Y \Rightarrow E(X) \geq E(Y).$$

Note: $E(Y) = a$, $P(X > a) \leq E(X)$

$$\Rightarrow P(X > a) \leq \frac{E(X)}{a}.$$  

Note: Markov inequality holds for any function of r.v. $X$, i.e., for any real valued function $g(X)$, the Markov's inequality is given by

$$P[g(X) > a] \leq \frac{E(g(X))}{a}, \quad a \neq 0.$$  

Proof: Let us define a function of r.v. $Y$, $g(Y)$

$$g(Y) = \begin{cases} a & \text{if } g(X) > a \\ 0 & \text{otherwise} \end{cases}$$

$$g(X) > g(Y)$$

$$\Rightarrow E(g(X)) \geq E(g(Y))$$

$$\Rightarrow E(g(Y)) = a \cdot P[g(X) > a] \leq E(g(X))$$

$$\Rightarrow P[g(X) > a] \leq \frac{E(g(X))}{a}, \quad a \neq 0.$$
**Problem 1.** If \( x \) be any r.v. such that \( M(t) = E(e^{tx}) \) exists for all \( t \), show that for any \( t > 0 \),
\[
P(tx > x^2 + \ln M(t)) < e^{-x^2}
\]

**Ans:**
We know that an exponential function is monotonically increasing.

So,
\[
P(tx > x^2 + \ln M(t))
= P(e^{tx} > e^{x^2 + \ln M(t)})
= P\left[e^{tx} > e^{x^2} \cdot e^{\ln M(t)}\right]
\]

Let \( g(x) = e^{tx} \) then by Markov's inequality, we have
\[
P(e^{tx} > e^{x^2} \cdot e^{\ln M(t)}) < \frac{E(e^{tx})}{e^{x^2} \cdot e^{\ln M(t)}} = \frac{M(t)}{M(t) \cdot e^{x^2}} = e^{-x^2} \quad (\text{Proved})
\]

**Problem 2.** For any random variable \( X \), show that,
\[
P[|X| > t] \leq \frac{1 + t^2}{t^2} E\left(\frac{X^2}{1 + X^2}\right) \quad \text{for any } t > 0.
\]

**Ans:**
Here,
\[
P[|X| > t]
= P\left[\frac{X^2}{1 + X^2} > \frac{t^2}{1 + t^2}\right]
= P\left[1 + X^2 > 1 + t^2\right]
= P\left[\frac{X^2}{1 + X^2} > \frac{t^2}{1 + t^2}\right]
\]

Now by Markov's inequality,
\[
P\left[\frac{X^2}{1 + X^2} > \frac{t^2}{1 + t^2}\right] \leq E\left(\frac{X^2}{1 + X^2}\right) \cdot \frac{1 + t^2}{1 + t^2} \quad (\text{Proved})
\]

**e.u. S.T.** \( P[X > t] < E(e^{ax})/e^{at} \).

**Ans:**
\[
P(ax > at) = P[X > t]
= P\left(e^{ax} > e^{at}\right) \quad [\because e \text{ is monotonically increasing}]
\]

By Markov's inequality,
\[
E(e^{ax}) \text{ exists when } a > 0.
\]
Problem 8. A fair die is rolled $n$ times. Find a lower bound to $n$ such that, the probability of at least one six in rolling is $\geq \frac{1}{2}$.

Ans: Let us define a random variable $X$ representing the number of six by throwing a die $n$ times.

\[ X \sim \text{bin} \left( n, \frac{1}{6} \right) \]

By Markov's inequality,

\[ P[X > 1] \leq \frac{E(X)}{1} \]

\[ \Rightarrow P[X > 1] \leq \frac{n}{6} \]

Again it is given that $P[X > 1] \geq \frac{1}{2}$

\[ \Rightarrow \frac{n}{6} \geq \frac{1}{2} \]

\[ \Rightarrow n \geq 3 \]

\[ \therefore \text{the die should be at least thrown 3 times.} \]

Problem 4. $X_1, X_2, \ldots, X_k$ are independent r.v.'s having zero mean and unit variance. Find an upper bound to,

\[ P \left[ \sum_{i=1}^{k} X_i^2 \geq \lambda k \right], \lambda > 0 \]

Ans: $X_1, X_2, \ldots, X_k$ are independent r.v.'s, with mean 0 and variance 1.

i.e., $E(X_i) = 0 \quad \forall \quad i = 1, 2, \ldots, k$

\[ V(X_i) = E(X_i^2) - E^2(X_i) \]

\[ \Rightarrow E(X_i^2) = 1 \]

\[ \sum_{i=1}^{k} E(X_i^2) = k \]

\[ \Rightarrow E \left( \sum_{i=1}^{k} X_i^2 \right) = k \]

\[ \therefore \text{the above holds good for independent variables} \]

Now by Markov's inequality,

\[ P \left[ \sum_{i=1}^{k} X_i^2 \geq \lambda k \right] \leq \frac{E \left( \sum_{i=1}^{k} X_i^2 \right)}{\lambda k} = \frac{k}{\lambda k} = \frac{1}{\lambda} \]

\[ \therefore \text{Required upper bound} = \frac{1}{\lambda}. \]
Chebyshev's Inequality:

Statement: For a random variable $X$ having finite mean and variance $\sigma^2$, then for any $t > 0$, the Chebyshev's inequality is given as follows:

$$P\left( |X - \mu| > t\sigma \right) \leq \frac{1}{t^2}$$

or

$$P\left( |X - \mu| \leq t\sigma \right) \geq 1 - \frac{1}{t^2}$$

Proof: In order to prove Chebyshev's inequality, we will first prove Markov's inequality.

Let us define a random variable $Z$,

$$Z = \begin{cases} a, & Y \geq a \\ 0, & \text{otherwise} \end{cases}$$

where $Y$ is another RV.

From the definition of $Z$, it is such that,

$$Y \geq Z \Rightarrow E(Y) \geq E(Z) \Rightarrow E(Y) \geq a \cdot P(Y \geq a) \Rightarrow P(Y \geq a) \leq \frac{E(Y)}{a}$$

This is the required Markov's inequality.

Now, for the RV $X$,

$$E(X) = \mu < \infty, \quad V(X) = \sigma^2 = E(X - \mu)^2 > 0$$

Now, $P\left( |X - \mu| > t\sigma \right) = P\left( (X - \mu)^2 > t^2 \sigma^2 \right)$

Now, let us choose $Y = (X - \mu)^2$ and $a = t^2 \sigma^2$, then by Markov's inequality, we have,

$$P\left( (X - \mu)^2 > t^2 \sigma^2 \right) \leq \frac{E(X - \mu)^2}{t^2 \sigma^2} = \frac{(x - \mu)^2}{t^2 \sigma^2}$$

$\therefore P\left( |X - \mu| > t\sigma \right) \leq \frac{1}{t^2}$

Hence proved.

1 - $<i>$ gives

$$P\left( |X - \mu| \leq t\sigma \right) \geq 1 - \frac{1}{t^2}$$

Hence proved.
The Equality Case: — Let us consider a r.v. with probability distribution,
\[ P[X = \mu + t\sigma] = \frac{1}{2t^2}, \text{ and} \]
\[ P[X = \mu] = 1 - \frac{1}{2t^2}. \]
where \( E(X) = \mu \), and \( \text{Var}(X) = \sigma^2 \).

If \( Y = |X - \mu| \), then
\[ P[Y = t\sigma] = \frac{1}{t^2} \quad \text{and} \quad P[Y = 0] = 1 - \frac{1}{t^2}. \]

Note,
\[ P[Y > t\sigma] = P[Y = t\sigma] = \frac{1}{t^2} \]
Therefore, \( P[|X - \mu| > t\sigma] \leq \frac{1}{t^2} \).
Hence equality holds for Chebyshev's inequality.

Another Proof: — Let \( Y = |X - \mu| \),
Then \( Y \) is a non-negative random variable with
\[ E(Y^2) = E((X - \mu)^2) = \sigma^2 \]
Now for any \( t > 0 \)
\[ P(Y > t\sigma) = P(Y^2 > t^2\sigma^2) \leq \frac{E(Y^2)}{t^2\sigma^2} \quad \text{[by Markov's inequality]} \]
\[ \leq \frac{\sigma^2}{t^2\sigma^2} = \frac{1}{t^2}. \]
Hence,
\[ P((X - \mu)^2 > t^2\sigma^2) \leq \frac{1}{t^2}. \]
\[ \Rightarrow P(|X - \mu| > t\sigma) \leq \frac{1}{t^2}. \]
Hence proved,
\[ 1 - \frac{1}{t^2} \]
gives
\[ P(|X - \mu| \leq t\sigma) \geq 1 - \frac{1}{t^2}. \]
Hence proved.
Problems 5. For a Laplace distribution with PDF:

\[ f(x) = \frac{1}{2} e^{-|x|}, \quad -\infty < x < \infty. \]

Find the minimum probability of an observation lying within the mean ± 3 s.d. interval.

\[ \text{OR} \]

Compare the value of \( P(|x-\mu| \leq 3\sigma) \) with the lower bound calculated by Chebyshev's inequality.

\[ \underline{\text{Ans.}}: \quad P(|x-\mu| \leq 3\sigma) \]

\[ = P(\mu-3\sigma \leq x \leq \mu+3\sigma) \]

\[ = \frac{1}{2} \int_{\mu-3\sigma}^{\mu+3\sigma} e^{-|x|} \, dx \]

\[ = \int_{\mu-3\sigma}^{\mu+3\sigma} e^{-x} \, dx \]

\[ = \left. -e^{-x} \right|_{0}^{\mu+3\sigma} \]

\[ = 1 - e^{-(\mu+3\sigma)} \]

\[ = 0.95 \quad \text{[} x \sim \text{Laplace}(0,1) \text{]} \]

By Chebyshev's inequality,

\[ P\left[ |x-\mu| \geq 3\sigma \right] \leq \frac{1}{92} \]

\[ \Rightarrow P\left[ |x-\mu| \leq 3\sigma \right] \geq 1 - \frac{1}{92} = \frac{91}{92} = 0.98 \]

Hence, the probability and the Chebyshev's upper bound is nearest to each other.

Problems 6. For the r.v. \( X \) having the following PDF:

\[ f(x) = e^{-x}, \quad x > 0 \]

\[ g(x) = \frac{x}{\lambda+1} \quad (0 < x < 2(\lambda+1)). \]

\[ \text{S.T.:} \quad P(0 < x < 2(\lambda+1)) \geq \frac{\lambda}{\lambda+1} \]

\[ \underline{\text{Ans.}}: \quad E(X) = (\lambda+1) = \frac{\lambda+1}{\lambda+1} \]

From Chebyshev's inequality,

\[ P\left[ \frac{|x-\mu|}{\sigma} < t \right] > 1 - \frac{1}{t^2} \]

\[ \Rightarrow P\left[ -\sigma t < (x-\mu) < \sigma t \right] > 1 - \frac{1}{t^2} \]

\[ \Rightarrow P\left[ -\frac{\lambda+1}{\lambda} < (x-\lambda) < \frac{\lambda+1}{\lambda} \right] > 1 - \frac{1}{\lambda+1} \]

\[ \Rightarrow P\left[ 0 < x < 2(\lambda+1) \right] > \frac{\lambda}{\lambda+1} \]

\[ \text{Design:} \quad t = 1 \]

\[ \Rightarrow \lambda = \frac{2+1}{2+1} \]
Problem 7. Let $X$ be an r.v. with mean $\mu$ and variance $\sigma^2 > 0$. If $\xi_a$ denotes the $a$th quantile of $X$, show that
\[
\mu - a \sqrt{\frac{1-a}{a}} \leq \xi_a \leq \mu + a \sqrt{1-a}.
\]

Ans. -
We know that $\xi_a$ satisfies the inequality $P(X \leq \xi_a) \geq a$.

If $\xi_a < \mu$, i.e., $\frac{\xi_a - \mu}{\sigma} < 0$, we have from Chebychev's inequality:
\[
a \leq P\left[ \frac{X - \mu}{\sigma} \leq \frac{\xi_a - \mu}{\sigma} \right] \leq \frac{1}{1 + \left( \frac{\xi_a - \mu}{\sigma} \right)^2}
\]
\[
\Rightarrow \frac{(\xi_a - \mu)^2}{\sigma^2} \leq \frac{1-a}{a}
\]
\[
\Rightarrow -\sqrt{\frac{1-a}{a}} \leq \frac{\xi_a - \mu}{\sigma} \leq \sqrt{\frac{1-a}{a}}.
\]

(Proved)

Problem 8. Let $f$ be a non-negative, non-decreasing function, prove that if $E(g(X-M))$ exists, where $M = E(X)$, then
\[
P[|X-M| > t] \leq \frac{E(g(|X-M|))}{g(t)}
\]

Ans. -
\[
P(g(|X-M|) > g(t)) \leq \frac{E(g(|X-M|))}{g(t)}
\]
But, $g(|X-M|) > g(t)$
\[
\Leftrightarrow |X-M| > t \quad \text{[for } g \text{ is non-decreasing & non-negative]}
\]
\[
\therefore P[|X-M| > t] \leq \frac{E(g(|X-M|))}{g(t)}. \quad \text{(Proved)}
\]
Problem 9. Let $F$ be the distribution function of the r.v. $X$ and $\mu$ and $\sigma^2 > 0$ are its mean and variance. Show that:

\[ F(x) \leq \frac{\sigma^2}{\sigma^2 + (x-\mu)^2} \quad \text{if } x \leq \mu \]

\[ F(x) \geq \frac{(x-\mu)^2}{\sigma^2 + (x-\mu)^2} \quad \text{if } x > \mu \]

**Ans:**

i) For $x \leq \mu$, let us take $x = \mu - t\sigma$, $t = \frac{x - \mu}{\sigma}$. Then one-sided Chebyshev's inequality gives:

\[ P[-X \geq -\mu - t\sigma] \leq \frac{1}{1 + t^2} \]

\[ \Rightarrow P[X \leq \mu - t\sigma] \leq \frac{1}{1 + t^2} \]

\[ \Rightarrow P[X \leq x] \leq \frac{1}{1 + \left(\frac{x - \mu}{\sigma}\right)^2} = \frac{\sigma^2}{\sigma^2 + (x-\mu)^2} \quad \text{for } x \leq \mu. \]

This result is trivially true for $x = \mu$, since in that case R.H.S $= 1$.

ii) For $x > \mu$, let us take $x = \mu + t\sigma$, and $t = \frac{x - \mu}{\sigma}$.

Then one-sided Chebyshev's inequality gives:

\[ P[-X \geq -\mu - t\sigma] \leq \frac{1}{1 + t^2} \]

\[ \Rightarrow P[X \geq x] \leq \frac{\sigma^2}{\sigma^2 + (x-\mu)^2} \quad \text{for } x > \mu. \]

This result is trivially true for $x = \mu$, since then the R.H.S becomes 0.
One sided Chebyshev's Inequality:

Statement: For an r.v. having finite mean $E(X) = \mu$ and finite variance $\sigma^2$, then we have

\[ P[X \geq \mu + \sigma t] \leq \frac{1}{1 + t^2} \quad \text{and} \]
\[ P[X \leq \mu - \sigma t] \leq \frac{1}{1 + t^2}. \]

Proof: Define an r.v. $Y = X - \mu$, $E(Y) = 0$ as $E(X) = \mu$.

\[ \therefore P(Y \geq \sigma t) \leq \frac{\sigma^2}{\sigma^2 + \sigma^2 t^2} = \frac{1}{1 + t^2} \quad (1) \]
\[ \therefore P(X \geq \mu + \sigma t) \leq \frac{1}{1 + t^2}. \]

\[ P(Y \leq -\sigma t) \leq \frac{\sigma^2}{\sigma^2 + \sigma^2 t^2} = \frac{1}{1 + t^2} \quad (2) \]
\[ \therefore P(X \leq \mu - \sigma t) \leq \frac{1}{1 + t^2}. \]

Lemma: If $X$ be an r.v. with mean zero and finite variance $\sigma^2$, then for any $a > 0$,

\[ i) P(X > a) \leq \frac{\sigma^2}{\sigma^2 + a^2} \quad \text{[Cc.U.1997]} \]
\[ ii) P(X \leq -a) \leq \frac{\sigma^2}{\sigma^2 + a^2}. \]

Proof: For any $b > 0$,

\[ P(X > a) = P(X + b > a + b) = P[(X + b)^2 > (a + b)^2] \leq \frac{E[(X + b)^2]}{(a + b)^2} \quad \text{[By Markov's Inequality]} \]
\[ \frac{E[X^2] + b^2}{(a + b)^2} = \frac{\sigma^2 + b^2}{(a + b)^2} = G_1(b), \quad \text{say}, \]
\[ G_1'(b) = \frac{(a + b)^2 \cdot 2b - (\sigma^2 + b^2)2(a + b)}{(a + b)^4} = 0 \]
\[ \Rightarrow ab = \sigma^2 \]
\[ \Rightarrow b = \frac{\sigma^2}{a} \Rightarrow b_{\min} = \frac{\sigma^2}{a}. \]

Thus putting $b_{\min} = \frac{\sigma^2}{a}$ in (1), we get,

\[ P(X > a) \leq \frac{\sigma^2 + \frac{\sigma^4}{a^2}}{(a + \frac{\sigma^2}{a})^2} = \frac{\sigma^2}{\sigma^2 + a^2}. \quad \text{[ii) no. proved]} \]
\[ X \text{ has mean } \mu, \text{ so } \mathbb{E}(X) = 0 \text{ and } \mathbb{V}(X) = \sigma^2 \]

\[ \therefore X = -X \]

\[ \Rightarrow P(-X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2} \]

\[ \Rightarrow P(X \leq -a) \leq \frac{\sigma^2}{\sigma^2 + a^2} \]

**Problem 10.** For the geometric distribution with PMF

\[ f(x) = \frac{1}{2^x}, \quad \text{if } x = 1, 2, 3, \ldots \]

Prove that

\[ P[|X - \mu| \leq 2] > \frac{1}{2} \]

**Ans:**

\[ E(X) = \sum_{x=1}^{\infty} x \frac{1}{2^x} \]

\[ = \frac{1}{2} \left[ 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots \right] \quad \Rightarrow \quad E(X) > \frac{1}{2} \]

\[ = \frac{1}{2} \left[ (1 + \frac{1}{2} + \frac{1}{2^2} + \ldots) + (\frac{1}{2} + \frac{1}{2^2} + \ldots) \right] \]

\[ \Rightarrow \quad E(X) = \frac{1}{2} \left[ 1 + \frac{1}{2} + \frac{1}{2^2} + \ldots \right] \]

\[ = \frac{1}{1 - \frac{1}{2}} = 2 \]

\[ \mathbb{V}(X) = \frac{\mathbb{V}(2X)}{(1/2)^2} = 2 \]

Putting \( \mu = 2 \), and \( \sigma = \sqrt{2} \)

\[ \therefore P[|X - \mu| \leq \sqrt{2}t] > 1 - \frac{1}{t^2} \quad \text{[By Chebyshev's inequality]} \]

Now, putting \( t = \sqrt{2} \),

\[ P[|X - \mu| \leq 2] > \frac{1}{2} \quad \text{(Proved)} \]
**SOME DISCRETE DISTRIBUTIONS**

**BERNOULLI'S DISTRIBUTION:**
X is said to be a Bernoulli random variable with parameter \( p, 0 < p < 1 \) if the p.m.f. of \( X \) is

\[
    f(x) = p[X = x] = p^x (1-p)^{1-x} \text{ I}_x (0, 1)
\]

i.e.

\[
    f(x) = \begin{cases} 
        p & \text{if } x = 1 \\
        1-p & \text{if } x = 0 
    \end{cases}
\]

**Bernoulli Trials:** A set of trials is said to be a set of Bernoulli trials if,

i) the result of any trials can be classified only under two categories, namely success and failure.

ii) the probability of success remains same in each trials.

iii) Trials are independent.

**BINOMIAL DISTRIBUTION:**
A discrete random variable \( X \) with mass points \( 0, 1, 2, \ldots, n \) is said to follow Binomial distribution with parameters \( n \) and \( p \) if its p.m.f. is given by,

\[
    f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x=0, 1, \ldots, n; \quad 0 < p < 1
\]

We write \( X \sim \text{Bin}(n, p) \).

**Result:** If \( X \) denotes the number of success in a set of \( n \) Bernoulli trials with probability of success \( p \) per trial, then \( X \sim \text{Bin}(n, p) \).

**Proof:** Let,

\( A_i \), the event that the \( i \)-th trial results in a success, \( i = 1, \ldots, n \).

The events \( A_1, A_2, \ldots, A_n \) are independent and with probability \( P(A_i) = p \), \( i = 1, \ldots, n \).

The mass points of \( X \) are \( 0, 1, 2, \ldots, n \). If \( x \) be any such mass points, then

\[
    P[X = x] = \binom{n}{x} P(A_1) P(A_2) \ldots P(A_{x-1}) P(A_{x+1}) P(A_{x+2}) \ldots P(A_n)
\]

\[
    = \binom{n}{x} p^x (1-p)^{n-x} \quad [\therefore A_i's \text{ are independent}]
\]

The p.m.f. of \( X \) is

\[
    f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \ldots, n; \quad 0 < p < 1
\]

i.e.

\[
    f(x) = \binom{n}{x} \sum_{x=0}^{n} \binom{n}{x} p^x (1-p)^{n-x}
\]
Problem 1. If $X \sim \text{Bin}(n,p)$, find the distribution of $Y = n - X$.

Since mass points of $X$ are 0, 1, ..., $n$.
Hence, the mass points of $Y$ are 0, 1, ..., $n$.

If $y$ be any such mass point then,

$$P(Y = y) = P(X = n - y) = P(X = n - y) = \binom{n}{y} p^y (1-p)^{n-y}$$

$$= \binom{n}{y} (1-p)^{n-y}$$

$X \sim \text{Bin}(n, 1-p)$

Problem 2. A box contains 2 coins with probability of heads 0.3 and 0.4 respectively. One of the coin is chosen at random and tossed $n$ times. Find the probability distribution of the number of heads obtain.

Let $X$ denotes the number of heads obtain.
Hence, the mass points of $X$ are 0, 1, 2, ..., $n$.

$A_i$ is the event that the $i$th coin is chosen. If $x$ be any mass point of $X$ then

$$P[X = x] = P[(X = x) \cap A_1] + P[(X = x) \cap A_2].$$

$$= P(A_1)P[(X = x) \mid A_1] + P(A_2)P[(X = x) \mid A_2].$$

$$= \frac{1}{2} \binom{n}{x} (0.3)^x (0.7)^{n-x} + \frac{1}{2} \binom{n}{x} (0.4)^x (0.6)^{n-x}$$

Factorial Moments:

Let $X \sim \text{Bin}(n, p)$, then the $r$th order

Factorial moment of $X$ is given by,

$$\mu_r' = E[(X)^r] = E\left[\prod_{i=1}^{r} X_i\right]$$

$$= \sum_{x=0}^{n} x^n \binom{n}{x} x_1 x_2 \cdots x_r$$

$$= \binom{n}{x} p^n x_1 x_2 \cdots x_r$$

$$= \binom{n}{r} p^n \sum_{x=0}^{n} \binom{n-x}{r} \frac{x^n}{x_1 x_2 \cdots x_r}$$

$$= \binom{n}{r} p^n \frac{(n-r)!}{(n-r)!} r^n$$

$$= \binom{n}{r} p^n (r + q)^{n-r}$$

$$= \binom{n}{r} p^n$$ for $0 \leq r \leq n$. 

$$= 0$$ otherwise.
Putting \( n = 1, \) we get —
\[
\mu_{[1]} = \mu_1 = E(X) = np
\]

Putting \( n = 2, \) we get —
\[
\mu_{[2]} = E[X(X-1)] = n(n-1)p^n
\]

\[
\therefore \text{ Var}(X) = E(X^2) - E(X)^2 = E[X(X-1)] + E(X) - E^2(X)
\]
\[
= n(n-1)p^n + np - np^n
\]
\[
= np(1-p) = npq.
\]

So, we can conclude that if \( X \sim \text{Bin}(n, p) \) then mean > Variance as \( p + q = 1, \) \( 0 < p < 1, \) i.e. \( np > npq. \)

\[\star 2.\] If \( X \sim \text{Bin}(n, p) \) then show that \( \text{Var}(X) \leq \frac{n}{4}. \)

\[
\text{Var}(X) = npq \leq np(1-p) \quad \text{differentiating w.r.t.} \quad p \quad \text{we get —}
\]
\[
\frac{d^2 \text{Var}(X)}{dp^2} = -2n < 0
\]
\[
\Rightarrow \quad \text{Var}(X) \text{ has a maxima at } p = \frac{1}{2}.
\]
\[
\therefore \quad \text{Var}(X) \text{ is maximum when } \text{Var}(X) = n \cdot \frac{1}{2} \left(1 - \frac{1}{2}\right) = \frac{n}{4}.
\]

\[\star 4.\] If \( X \sim \text{Bin}(n, p), \) find \( E\left(\frac{1}{x+1}\right) \) & \( E\left[\frac{1}{(x+1)(x+2)}\right] \)

\[
E\left(\frac{1}{x+1}\right) = \sum_{x=0}^{n} \frac{1}{1+x} \left(\begin{array}{c} n \\ x \end{array}\right) p^x q^{n-x}
\]
\[
= \frac{1}{P(n+1)} \sum_{x=0}^{n} \left(\begin{array}{c} n+1 \\ x+1 \end{array}\right) p^{x+1} q^{n-x-1}
\]
\[
= \frac{1}{P(n+1)} \sum_{x=0}^{n} \left(\begin{array}{c} n+1 \\ x \end{array}\right) p^x q^{n+1-x} - \left(\begin{array}{c} n+1 \\ 0 \end{array}\right) p^n q^{n+1}
\]
\[
= \frac{1}{P(n+1)} \left[ (1 - q^{n+1}) \right]
\]
Coefficient of Skewness:

\[ \beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{n^p q^n (1-2p)^n}{npq} \left(1-2p\right) \leq 0 \text{ according to } \mu_2 \]

\[ \Rightarrow \text{ The distribution is symmetric if } (1-2p)=0 \Rightarrow \mu_2^{3/2} = \frac{1}{2} \text{ or } p=\frac{1}{2} \]

\[ +\text{, very skewed if } (1-2p)>0 \Rightarrow \mu_2^{3/2} < \frac{1}{2} \text{ or } p<\frac{1}{2} \]

\[ -\text{, very skewed if } (1-2p)<0 \Rightarrow \mu_2^{3/2} > \frac{1}{2} \text{ or } p>\frac{1}{2} \]

Intuitively, if \( p=\frac{1}{2} \), \( P(X=x) = \left(\frac{n}{2}\right)^{\frac{n}{2}} = \left(\frac{n-x}{2}\right)^{\frac{n}{2}} \)

\[ \Rightarrow \text{ The resulting distribution is symmetric.} \]

\[ p=a \quad p>a \quad p<q \]

Coefficient of Kurtosis:

\[ \beta_2 = \frac{\mu_4}{\mu_2^2} - 3 + \frac{(1-Gpa)}{npq} \]

Now, \( \beta_2 \leq 3 \) according as \( (1-Gpa) \leq 0 \), i.e., \( p \leq \frac{1}{2} \).

\[ \Rightarrow \text{ The distribution is mesokurtic when } (1-Gpa) = 0 \text{, i.e., } p = \frac{1}{2} \]

\[ \Rightarrow \text{ leptokurtic when } (1-Gpa) > 0 \text{, i.e., } p < \frac{1}{2} \]

\[ \Rightarrow \text{ platykurtic when } (1-Gpa) < 0 \text{, i.e., } p > \frac{1}{2} \]

\[ \star 5. \text{ If } X \sim \text{Bin}(n,p), \text{ show that } \Rightarrow \text{ Cov} \left(\frac{X}{n}, \frac{n-X}{n}\right) = \frac{p - p^2}{n} \]

\[ \text{i) } E \left(\frac{X}{n} - p\right)^2 = \frac{p - p^2}{n} \]

\[ \text{ii) } \text{ Cov} \left(\frac{X}{n}, \frac{n-X}{n}\right) = \frac{p - p^2}{n} \]

\[ \star 6. \text{ If } X \& Y \text{ are independent, Bin}(n,p) \text{. Then find the distribution of } (X-Y). \]

Mass points of \( X \) are 0, 1, 2, ..., \( n \) & same for \( Y \).

Then the mass points of \( (X-Y) \) are \(-n, -n+1, -n+2, ..., 0, 1, 2, ..., n\).

\[ P(X-Y=\pm) = \sum_{k=0}^{n} P[X=k, Y=k-\pm] \]

\[ = \sum_{k=0}^{n} P[X=k] P[Y=k-\pm] \quad [\Rightarrow X \& Y \text{ are independent}] \]

\[ = \sum_{k=0}^{n} \binom{n}{k} p^k q^{n-k} \binom{n-k}{k-\pm} p^{k-\pm} q^{n-k-\pm} \]

\[ = \sum_{k=0}^{n} \binom{n}{k} \binom{n-k}{k-\pm} p^{2k-\pm} q^{2n-2k+\pm} \quad \Rightarrow \pm = -n \text{ or } n \]
Recurrence Relation Regarding Probabilities:

1. If $X \sim \text{Bin}(n, p)$

   
   $P(X = \alpha) = \binom{n}{\alpha} p^\alpha q^{n-\alpha}; \quad 0 < p < 1$

   
   $P(X = \alpha - 1) = \binom{n}{\alpha - 1} p^{\alpha - 1} q^{n-\alpha+1}$

   
   So,

   
   $\frac{P(X = \alpha)}{P(X = \alpha - 1)} = \frac{\binom{n}{\alpha}}{\binom{n}{\alpha - 1}} \cdot \frac{\frac{p}{q}}{\frac{\alpha}{\alpha - 1}} = \frac{n - \alpha + 1}{\alpha} \cdot \frac{p}{q}$

   
   

   

   $\therefore P(X = \alpha) = \frac{(n - \alpha + 1)p}{\alpha q} \cdot P(X = \alpha - 1)$

2. If $P[X = \alpha] = P_{\alpha}$ for a binomially distributed random variable, say, with parameters $(n, p)$, then show that

   
   $\frac{P_1}{P_0} > \frac{P_2}{P_1} > \frac{P_3}{P_2} > \cdots > \frac{P_n}{P_{n-1}}$.

   
   From the above (1), we get,

   
   $\frac{P(X = \alpha)}{P(X = \alpha - 1)} = \frac{n - \alpha + 1}{\alpha} \cdot \frac{p}{q}$

   
   

   

   $\therefore \frac{P(X = \alpha)}{P(X = \alpha - 1)}$ is a decreasing function of $\alpha$.

   

   $\therefore \frac{P_1}{P_0} > \frac{P_2}{P_1} > \cdots > \frac{P_n}{P_{n-1}}$.

3. Show Binomial variate as a sum of independent Bernoulli variates.

   Suppose, $X_1, X_2, \ldots, X_n$ are independently distributed Bernoulli random variables with parameter $p$.

   i.e., $p$.$g$.$f$. of $X_i$ is

   
   $P(X_i = 1) = p; \quad P(X_i = 0) = q = 1 - p, \quad \forall \alpha = 0, 1$.

   We know $p$.$g$.$f$. of $S_n = \sum_{i=1}^{n} X_i$

   
   $P_{S_n}(t) = \left[ P_{X_i}(t) \right]^n$

   
   $\therefore P_{X_i}(t) = E(t^{X_i}) = t^0(1-p) + tp$

   
   

   

   

   $= 1 - p + tp$

   

   

   $P_{S_n}(t) = (1 - p + tp)^n$

   

   

   $= \sum_{\alpha=0}^{n} \binom{n}{\alpha} (pt)^\alpha (1-p)^{n-\alpha}$

   

   $\therefore [S_n = \alpha] =$ coefficient of $t^\alpha$ in the expansion of $P_{S_n}(t)$.

   

   $= \binom{n}{\alpha} p^\alpha (1-p)^{n-\alpha}; \quad \alpha = 0, 1, \ldots, n.$
Mode of the Binomial Distribution

Let \( X \sim \text{Bin}(n, p) \), the p.m.f. of \( X \) is

\[
f(x) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, \ldots, n \]

where \( 0 < p < 1 \) and \( p + q = 1 \).

The mode of the distribution is the value \( x \) of \( x \) for which \( f(x) \) is maximum. [mode = most probable value]

For \( x = 1, 2, \ldots, n \),

\[
\frac{f(x)}{f(x-1)} = \frac{\binom{n}{x} p^x q^{n-x}}{\binom{n}{x-1} p^{x-1} q^{n-x+1}} = \frac{(n-x+1)p}{xq}
\]

Now,

\[
f(x) \geq f(x-1) \quad \text{according as} \quad \frac{(n-x+1)p}{xq} \geq 1 \]

\[
\Rightarrow np - xp + p \geq x - xp \quad \Rightarrow (n+1)p \geq x
\]

\[
\Rightarrow f(x) \geq f(x-1) \quad \text{according as} \quad x \leq (n+1)p
\]

\[
\text{Case-1:} \quad (n+1)p = \text{integer} = k \quad (\text{say})
\]

Now, for \( x = k \),

\[
\Rightarrow f(k) = f(k-1) \quad \ldots \ldots (1)
\]

For \( x < k \)

\[
\Rightarrow f(x) > f(x-1) \quad \text{for} \quad x = 1, \ldots, k-1
\]

\[
\Rightarrow f(0) < f(1) < \ldots \ldots < f(k-1) \quad \ldots \ldots (2)
\]

Again, for \( x > k \)

\[
\Rightarrow f(x) < f(x-1) \quad \text{for} \quad x = k+1, k+2, \ldots, n
\]

\[
\Rightarrow f(k) > f(k+1) > \ldots \ldots > f(n) \quad \ldots \ldots (3)
\]

Combining \( (1), (2), (3) \), we get,

\[
f(0) < f(1) < \ldots \ldots < f(k-1) = f(k) > f(k+1) > \ldots \ldots > f(n)
\]

\[
\Rightarrow f(x) \text{ is maximum at } x = k \text{ and } x = k-1
\]

\[
\Rightarrow \text{the modes of the distribution are at } x = (n+1)p \text{ and } (n+1)p
\]
Case-2: \( (n+1)p \) is not an integer, let \( m = \lfloor (n+1)p \rfloor = \text{largest integer obtained in } (n+1)p \).

Obviously, \( m < (n+1)p < m+1 \).

Now, \( x = m \)
\( \Rightarrow x < (n+1)p \)
\( \Rightarrow f(x) > f(x-1) \) for \( x = 1, \ldots, m-1 \)
\( \Rightarrow f_0 < f(1) < \ldots < f(m-1) \) \( \cdots \) \( \cdots \) \( \cdots \) \( \cdots \) \( \cdots \)
\( x > m \)
\( \Rightarrow x > m+1 > (n+1)p \)
\( \Rightarrow x > (n+1)p \)
\( \Rightarrow f(x) < f(x-1) \) for \( x = m+1, \ldots, n \)
\( \Rightarrow f(m) > f(m+1) > \ldots > f(n) \)

Combining \( 4, 5, 6 \); we get
\( f(0) < f(1) < f(2) < \ldots < f(m) > f(m+1) > \ldots > f(n) \)
\( \therefore f(x) \) is maximum at \( x = m \)
\( \therefore \) the mode of the distribution are at \( x = (n+1)p \).

Example: \( i) X \sim \text{Bin}(12; \frac{1}{3}) \); \( ii) X \sim \text{Bin}(11; \frac{1}{3}) \)

\( \therefore \) Mode = \( \left[ \frac{12}{3} \right] = 4 \)

9. If \( X \sim \text{Bin}(m; p) \), find the restricted range of \( p \) for which there will be a single mode \( x = 1 \), where \( m = 21 \).

\( \therefore (n+1)p = (21+1)p = 22p \)

It is given that \( \lfloor 22p \rfloor = 1 \), hence \( m = 1 \) \( \therefore \) by the property of the binomial function
\( 1 \leq 22p < 2 \)
\( \Rightarrow \frac{1}{22} \leq p < \frac{1}{11} \)
10. If \( X \sim \text{Bin}(28, \frac{1}{2}) \), \( s \) = integer, show that,

i) the most probable value of \( X \) is 8.

\[
p[X = 8] = \frac{(28-1)(28-2)}{28} = 4.2
\]

ii) \[
\frac{1}{\sqrt{28}} < p[X = 8] < \frac{1}{\sqrt{28+1}} \quad \text{[Wesu'11]}
\]

iii) \[
\frac{(n+1)p}{2} = (28+1) \frac{1}{2} = 28.5
\]

\[
\Rightarrow [n+1]P = [28+\frac{1}{2}] = 28.5, \quad \text{Hence the proof.}
\]

\[
P[X = 8] = \binom{n}{8} p^8 (1-p)^{n-8} = \frac{28!}{8! 20!}
\]

\[
= \frac{1}{2^{28}} \cdot \frac{28!}{8! 20!} = \frac{28!}{2^{28} 8! 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!}
\]

\[
= \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!}
\]

\[
= \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!}
\]

\[
= \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!} = \frac{28!}{2^{28} \cdot 8! \cdot 20!}
\]

\[
\Rightarrow P[X = 8] < \frac{1}{p[X = 8] (28+1)}
\]

\[
\Rightarrow P[X = 8] < \frac{1}{\sqrt{28+1}} \quad \text{10}
\]
No we, \( \frac{2(k-1)}{2k-1} < \frac{2k-1}{2k} \)

Putting \( k = 2, 3, \ldots, 8 \); successively we get:

\[
\frac{2}{3} < \frac{3}{4} < \frac{4}{5} < \frac{5}{6} < \frac{6}{7} < \frac{7}{8}
\]

\[
\frac{2(8-1)}{28-1} < \frac{28-1}{28}
\]

Multiplying we get,

\[
\frac{2, 4, 6, \ldots, (28-2)}{3, 5, 7, \ldots, (28-1)} < \frac{1, 3, 5, \ldots, (28-1)}{4, 6, \ldots, 28}
\]

\[
\Rightarrow \frac{1}{2 \cdot 8 \cdot P[X=8]} < 2P[X=8] \Rightarrow P[X=8] > \frac{1}{2 \cdot 8}
\]

\[
\Rightarrow P[X=8] > \frac{1}{2 \sqrt{8}}
\]

\[
\Rightarrow \frac{1}{2 \sqrt{8}} < P[X=8] < \frac{1}{\sqrt{28+1}}
\]

\[
\star \text{II. If } X \sim \text{Bin}(n, p) \text{. Show that } f(k) < \frac{1}{k-np} \text{ for } k > np.
\]

For, \( k > np \),

\[
\sum_{x=\lceil np \rceil+1}^{k} f(x) < 1 \quad \text{[since it's a probability]}
\]

since \( f(x) \) is a decreasing function of \( x \), where \( x > (n+1)p \).

\[
f(x) = f(x-1) \text{ when } x = (n+1)p
\]

\[
(n+1)p \leq \lceil np \rceil + 1
\]

then

\[
1 > \sum_{x=\lceil np \rceil+1}^{k} f(x) > \sum_{x=\lceil np \rceil+1}^{k} f(k)
\]

\[
= f(k) \left[ k - \lceil np \rceil \right]
\]

\[
\Rightarrow f(k) < \frac{1}{k-\lceil np \rceil} < \frac{1}{k-np} \text{ when } k > np
\]
Mean Deviation about Mean:

\[ MD_\mu(x) = E|X - \mu| \]
\[ = \sum_{x=0}^{n} |x - \mu| f(x) \]
\[ = \sum_{x=0}^{\mu} (\mu - x) f(x) + \sum_{x=\mu+1}^{n} (x - \mu) f(x) \]
\[ = 2 \sum_{x=\mu+1}^{n} (x - \mu) f(x) \quad \text{[since } E(X-\mu) = 0] \]
\[ = 2 \sum_{x=\mu+1}^{n} (x - \mu) f(x) \]
\[ = 2 \sum_{x=\mu+1}^{n} (x - \mu) P a^{n-x} \]
\[ = 2 \left[ \sum_{x=m_0}^{n} \alpha_{n-1} P a^{n-x} - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]
\[ = 2n_p \left[ \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} (1 - P) - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]
\[ = 2n_p \left[ \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]
\[ = 2n_p \left[ \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]
\[ = 2n_p \left[ \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]
\[ = 2n_p \left[ \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} - \sum_{x=m_0}^{n-1} \alpha_{n-1} P a^{n-x} \right] \]

Using Stirling's approximation formula:

\[ n! \approx \sqrt{2\pi n} e^{-n} n^{n+\frac{1}{2}} \]

\[ MD_\mu(x) = \frac{2m_0 q \sqrt{2\pi} e^{-m_0} m_0^{m_0+\frac{1}{2}}}{\sqrt{2\pi} e^{-m_0} m_0^{m_0+\frac{1}{2}} (n-m_0)^{n-m_0+\frac{1}{2}}} \]

For large \( n \):

\[ \frac{m_0}{n} \to \mu \quad \text{and} \quad \frac{n-m_0}{n} \to 0 \quad \frac{mp}{m_0} \to 1 \]

\[ E|X - \mu| = \sqrt{\frac{2n_p q}{\pi}} \quad \text{[Gazavi's Ratio]} \]
Distribution function in terms of Incomplete Beta function:

\[ I_x(m,n) = \int_0^x u^{m-1}(1-u)^{n-1} du / B(m,n) = \int_0^x u^{m-1}(1-u)^{n-1} du / B(m,n) \quad ; \quad m>0, n>0 \]

P.d.f. of Beta distribution of 1st kind is

\[ f(x) = \frac{x^{m-1}(1-x)^{n-1}}{B(m,n)} \quad \text{if} \quad 0<x<1 \]

Let, \( X \sim \text{Bin}(n,m) \) then show that,

\[ P[X \leq k] = \sum_{j=0}^k \binom{n}{j} p^j q^{n-j} \]

**Proof:**

\[
\begin{align*}
P[X \leq k] &= \sum_{j=0}^k \binom{n}{j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{n-k}} \sum_{j=0}^k \frac{k!(n-k-j)!}{j!(n-j)!} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \frac{(k-j)!}{(n-j)!} \binom{n-j}{j-k} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \frac{\binom{k}{j} \binom{n-k}{n-j}}{\binom{n}{n-j}} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{k}{j} \binom{n-k}{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\
&= \frac{1}{\binom{n}{k}} \sum_{j=0}^k \binom{n-k}{n-j} \binom{k}{j} (1-q)^{n-j} p^j q^{n-j} \\& From the above relationship, it could be said that, \\
I_q(n-k,k+1) = 1 - I_p (k+1, n-k).
Problem No. 17. A drunk performs a random walk, over position \( 0, \pm 1, \pm 2, \ldots \), as follows:

He starts at 0. He takes successive one-unit steps, going to the right with probability \( p \) and to the left with probability \( (1-p) \). His steps are independent. Let \( X \) denote his position after \( n \) steps. Find the distribution of \( \frac{X+n}{2} \) and then find \( E(X) \).

Solu.: \( X_i = \begin{cases} -1 & \text{if the steps to the left with probability } (1-p) \\ 1 & \text{if the steps to the right with probability } p \end{cases} \)

for \( i = 1, 2, \ldots, n \).

Then, \( \frac{X_i + 1}{2} = \begin{cases} 0 & \text{if the steps to the left with probability } (1-p) \\ 1 & \text{if the steps to the right with probability } p \end{cases} \)

is a Bernoulli random variable, \( i = 1, 2, \ldots, n \).

Here, \( X_i \)’s are i.i.d. Bernoulli r.v.’s, \( i = 1, 2, \ldots, n \), then,

\[
\sum_{i=1}^{n} \frac{X_i + 1}{2} \sim \text{Bin}(n, p)
\]

\[
\Rightarrow \frac{1}{2} \sum_{i=1}^{n} X_i + n \sim \text{Bin}(n, p) \Rightarrow \frac{X+n}{2} \sim \text{Bin}(n, p)
\]

Here, \( E\left(\frac{X+n}{2}\right) = np \)
\( \Rightarrow E(X) + n = 2np \)
\( \Rightarrow E(X) = (2p-1)n \)

Problem 18. Let mutually independent r.v.’s \( X_1, X_2, X_3 \) have the same p.d.f. \( f(x) = \begin{cases} 8x^2 & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \)

Find the probability that exactly 2 of the 3 variables exceed \( \frac{1}{2} \).

Solu.: consider \( X_i > \frac{1}{2} \) as success, \( i = 1, 2, 3 \). \( X_i \)'s are independent Bernoulli trials with \( p = \frac{3}{2} \).

Define, \( Y = \) the no. of \( X_i \)'s which are \( > \frac{1}{2} \)
\( = \) the no. of successes in 3 Bernoulli trials,

\( \sim \text{Bin}(n=3, p = \frac{7}{8}) \)

Required prob. = \( P[Y=2] = \binom{3}{2} p^2 (1-p)^{3-2} = \left(\frac{3}{2}\right) \left(\frac{7}{8}\right)^2 \left(\frac{1}{8}\right) \)

\( = \frac{21}{8} \times \frac{7}{8} = \frac{147}{512} \)
Negative Binomial Distribution:

Negative Binomial distribution will be appropriate if the counts of an event which occurs in clusters such as,

1. The distribution of mosquito bites,
2. The distribution of the number of eggs laid by an insect on leaves of a tree.

Definition: \( X \) is said to have a negative binomial distribution (or inverse binomial distribution or waiting-time binomial distribution) if the p.m.f. of \( X \) is of the form

\[
P[X = k] = \binom{k + \alpha - 1}{k} p^{\alpha} q^{k} I_{k \geq 0, \alpha > 0, p \in [0, 1]}
\]

We denote, \( X \sim NB(\alpha, p) \)

\[
\binom{k + \alpha - 1}{k} = \frac{(k + \alpha - 1)!}{k!}
\]

\[
= \frac{(-\alpha)(-\alpha-1) \cdots (-\alpha-k+1)}{\alpha!}
\]

\[
= \frac{(-1)^k (\alpha + \alpha - 1) \cdots (\alpha + k - 1)}{\alpha!}
\]

\[
= \frac{(-1)^k (-\alpha)(-\alpha-1) \cdots (-\alpha-k+1)}{\alpha!}
\]

\[
= (-1)^\alpha \binom{-\alpha}{k}
\]

\[
= (-1)^\alpha \binom{-\alpha-1}{k}
\]

\[
i.e., P[X = k] = \binom{k + \alpha - 1}{k} p^{\alpha} q^{k} I_{k \geq 0, \alpha > 0, p \in [0, 1]}
\]

We know,

\[
p^{\alpha} (1-q)^{-\alpha} = 1
\]

\[
\sum_{k=0}^{\infty} \binom{k + \alpha - 1}{k} p^{\alpha} q^{k} = 1,
\]

i.e.,

\[
\sum_{k=0}^{\infty} \binom{k + \alpha - 1}{k} p^{\alpha} q^{k} = 1
\]

So, it's a p.m.f.

The masses of negative binomial distribution over the different points are obtained as the different terms in the expansion of

\[
p^{\alpha} (1-q)^{-\alpha} = \left(\frac{p}{p - q}\right)^{-\alpha}
\]

Here as the index is a negative integer, so the distribution is named as negative binomial distribution.
The distribution is also called a Waiting Time binomial distribution because the mass points of $x$ can be obtained from a sequence of Bernoulli trials with a stopping rule:

Consider a sequence of Bernoulli trials with success probability $p$. Let the trials be repeated until the occurrence of $r$ successes.

Define $X = \text{No. of failures preceding the } r\text{th success.}$

$p[X=x] = p[\text{first } (x+r-1)\text{ trials result in } (x-1)\text{ successes and } (x+r)\text{ th trial results in a success}]$

$$= \binom{x+r-1}{x-1} p^{x-1} q^r p$$

Since the event $\{X=x\}$ may occur in $\frac{(x+r-1)!}{x!(r-1)!}$ mutually exclusive and exhaustive and equally probable ways.

**Probability Model:** Consider an indefinite series of Bernoulli trials. Let $p (0 < p < 1)$ and the probability of a success in a trial. Let the trials be repeated until we are getting $r$ success column $r$ is a pre-fixed positive integer. Let,

$Y = \text{No. of trials required to get } r\text{ successes.}$

Mass points of $Y$ are $r, r+1, r+2, \ldots \infty$. If $y$ be any such mass point then,

$p[Y=y] = p[\text{exactly } y\text{ trials required to get } r\text{ success}]$

$$= p[\text{In first } (y-1)\text{ trials we get } (y-1)\text{ successes and } y\text{ th trial results a success}]$$

$$= \binom{y-1}{y-1} p^{y-1} q^{r-y} p$$

$$= \binom{y-1}{y-1} p^{y-1} q^{r-y} I_y \{y, r+1, \ldots \}$$

*OR*

$p[\text{a particular case}]$

$= p[SSF \ldots \ldots SFFF \ldots FS]$

The particular case occurs $\binom{y-1}{r-1}$ times.

$$p[Y=y] = \binom{y-1}{r-1} p^{y-1} q^{r-y} p$$

$$= \binom{y-1}{r-1} p^{y-1} q^{r-y} p$$
Alternatively, if $X$ denotes the number of failures preceding the $n$th success.

\[
\therefore Y = n + X \implies X = Y - n
\]

mass points of $X$ are $0, 1, 2, \ldots \infty$.

If $x$ be any such mass points then

\[
P[X = x] = P[Y - n = x] = P[Y = n + x] = \binom{n + x - 1}{x} p^n q^x
\]

\[
= \binom{n}{x} (-1)^x p^n q^x
\]

\[
= (-1)^x p^n (-q)^x
\]

Factorial Moments:

Let $X \sim NB(n, p)$.

The $k$th order factorial moment is

\[
\mu [E_k] = E [X^k] = \sum_{x = 0}^{\infty} (x)^k f(x) = \sum_{x = 0}^{\infty} \binom{x}{k} (x + n - 1) p^n q^x
\]

\[
= \sum_{x = k}^{\infty} \binom{x}{k} \frac{(x + n - 1)!}{(x - k)! (n - 1)!} p^n q^x
\]

\[
= q^k \frac{(k + n - 1)!}{(n - 1)!} \sum_{x = k}^{\infty} \binom{x}{k} \frac{(x - k)! (x + k - 1)! p^n q^{x-k}}{y! (n + k - y)!}
\]

\[
= q^k \frac{(k + n - 1)!}{n!} \sum_{j = 0}^{\infty} \frac{(n + y + k - 1)! p^n q^y}{j! (n + k - y)!}
\]

\[
= q^k \frac{(k + n - 1)!}{n!} \sum_{j = 0}^{\infty} \frac{(n + y + k - 1)! p^n q^y}{j! (n + k - y)!}
\]

\[
= \mu [E_{k+1}] = \binom{n}{k} \left( \frac{q}{p} \right)^k
\]

\[
\mu [E_{k+1}] = \binom{n}{k} \left( \frac{q}{p} \right)^k
\]

\[
= p^k \cdot \frac{q}{p}
\]

\[
\mu [E_{2}] = \binom{n}{2} \left( \frac{q}{p} \right)^2 = \binom{n+1}{2} \left( \frac{q}{p} \right)^2
\]

\[
\mu [E_{2}] = \binom{n}{2} \left( \frac{q}{p} \right)^2 = \binom{n+1}{2} \left( \frac{q}{p} \right)^2
\]

\[
\mu_2 = \mu [E_{2}] - \mu [E_{1}] + \mu [E_{0}]
\]

\[
= n \binom{n+1}{2} \left( \frac{q}{p} \right)^2 - \binom{n}{1} \frac{q}{p} + n \frac{q}{p}
\]

\[
= n \binom{n+1}{2} \left( \frac{q}{p} \right)^2 - \binom{n}{1} \frac{q}{p} + n \frac{q}{p}
\]

\[
\therefore \text{mean < variance.}
\]
4) \( X \) = Number of failures preceding \( k \)th success in a sequence of Bernoulli trials with success probability \( p \).

\[
P[X \leq k] = P[X+k \leq k+1] = P[N \leq k+1] = P[Z \geq k+1]
\]

\( N = X+k \) = Number of trials required to get \( k \) successes.

\( Z \) = Number of successes out of \((k+1)\) trials, and \( Z \sim \text{Bin}(k+1, p) \).

Recursion Relation for Central Moments:

If \( X \sim \text{NB}(n, p) \), then

\[
\mu_k = E[(X - E(X))^k] = \sum_{x=0}^{\infty} \left[ x - \frac{n}{p} \right]^{k} \left( \frac{n}{p} \right)^{x} \left( 1 - \frac{n}{p} \right)^{n-x}
\]

Differentiating both sides w.r.t. \( p \), we get:

\[
\frac{d\mu_k}{dp} = \sum_{x=0}^{\infty} \left[ x - \frac{n}{p} \right]^{k+1} \left( \frac{n}{p} \right)^{x+1} \left( 1 - \frac{n}{p} \right)^{n-x-1}
\]

\[
= \sum_{x=0}^{\infty} \left( x + 1 \right) \left( x - \frac{n}{p} \right)^{k} \frac{n}{p} \left( 1 - \frac{n}{p} \right)^{n-x-1}
\]

\[
= \frac{n}{p} \mu_k + \sum_{x=0}^{\infty} \left( x + 1 \right) \left( x - \frac{n}{p} \right)^{k+1} \frac{n}{p} \left( 1 - \frac{n}{p} \right)^{n-x-1}
\]

\[
= \frac{n}{p} \mu_k + \frac{n}{p} \mu_{k+1}
\]

\[
\mu_{k+1} = \frac{n}{p} \mu_k + \frac{n}{p} \mu_{k+1}
\]

Put, \( k = 1 \), \( \mu_0 = 1, \mu_1 = 0 \),

\[
\mu_2 = \frac{n}{p} \left[ \frac{n}{p} \right]^2 = \frac{n^2}{p^2}
\]

Putting \( k = 2 \),

\[
\mu_3 = \frac{n}{p} \left[ \frac{2n+1}{p^3} \right] = \frac{n^2}{p^3} (2n+1)
\]
Distribution function in terms of Incomplete Beta function.

Result: If \( X \sim NB(\alpha, p) \) then for any non-negative integer \( k \),

\[
P[X \leq k] = I_p(\alpha, k+1),\quad \text{where}
\]

\[
I_s(a, b) = \frac{1}{B(a, b)} \int_0^s t^{a-1} (1-t)^{b-1} dt
\]

Proof:

\[
P[X \leq k] = \sum_{x=0}^{k} \binom{n+x-1}{x} p^n q^x
\]

\[
= \frac{1}{B(\alpha, k+1)} \sum_{x=0}^{k} \frac{\Gamma(\alpha) \Gamma(k+1)}{\Gamma(\alpha+k+1)} \binom{n+x-1}{x} p^n q^x
\]

\[
= \frac{1}{B(\alpha, k+1)} \sum_{x=0}^{k} \binom{k}{x} \frac{(n+x)! (k)!}{(n+k)! (k-x)! x!} p^n q^x
\]

\[
= \frac{1}{B(\alpha, k+1)} \sum_{x=0}^{k} \binom{k}{x} \frac{\Gamma(k-x+1) \Gamma(n+x)}{\Gamma(n-k+1)} p^n q^x
\]

\[
= \frac{1}{B(\alpha, k+1)} \sum_{x=0}^{k} \binom{k}{x} B(n+x, k-x+1) p^n q^x
\]

\[
= \frac{1}{B(\alpha, k+1)} \int_p^0 u^{n-x-1} \sum_{x=0}^{k} \binom{k}{x} (qu)^x (1-u)^{k-x} du
\]

\[
= \frac{1}{B(\alpha, k+1)} \int_p^0 u^{n-1} (qu+1-u)^k du
\]

\[
= \frac{1}{B(\alpha, k+1)} \int_{pu}^0 (1-pu)^{k-1} (pu)^{n-1} p du
\]

\[
= \frac{1}{B(\alpha, k+1)} \int_0^1 t^{n-1} (1-t)^{k-1} dt
\]

\[
= I_p(\alpha, k+1).
\]
Mean Deviation About Mean:

\[ \text{List } X \sim NB(\alpha, p) \]

\[
E(X) = \mu = \frac{\alpha p}{1 - (1 - p)^\alpha} \]

\[
\text{MD}_\mu(X) = E |X - \mu| = \sum_{x=0}^{\infty} |x - \frac{\alpha p}{1 - (1 - p)^\alpha}| f(x)
\]

\[
= 2 \sum_{x=0}^{\infty} (\frac{\alpha p}{1 - (1 - p)^\alpha} - x) f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} (\alpha x - (1 - p)^\alpha) f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} [(x + \alpha) q - x] f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} (x + \alpha) q f(x) - \frac{2}{\alpha} \sum_{x=0}^{\infty} x f(x)
\]

\[
\text{Let, } g(x) = x f(x) = x \cdot \left(\frac{\alpha x - (1 - p)^\alpha}{\alpha x - (1 - p)^\alpha}\right) p^n q^{x-1}
\]

\[
= \left(\frac{\alpha x + (1 - p)^\alpha}{\alpha x - (1 - p)^\alpha}\right) p^n q^x = q \cdot \left(\frac{x + \alpha}{x - \alpha}\right) f(x)
\]

\[
\text{MD}_\mu(x) = \frac{2}{\alpha} \sum_{x=0}^{\infty} \left[ g(x+1) - g(x) \right] f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} \left[ \left(\frac{x + \alpha + 1}{x - \alpha}\right) - \left(\frac{x + \alpha}{x - \alpha}\right) \right] f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} \left[ \frac{x + \alpha + 1}{x - \alpha} \right] f(x)
\]

\[
= \frac{2}{\alpha} \sum_{x=0}^{\infty} \left[ \frac{x + \alpha + 1}{x - \alpha} \right] f(x)
\]

\[
\text{Now, } P[X = \mu_0] = \left(\frac{\mu_0 + n - 1}{\mu_0}\right) \cdot p^n q^{\mu_0}
\]

\[
= \left(\frac{\mu_0 + n - 1}{\mu_0}\right) \cdot p^n q^{\mu_0}
\]

\[
\approx \sqrt{\frac{2\pi}{2n}} \cdot e^{-n} \cdot n^{\frac{1}{2}} \text{ for large } n, \text{ by Stirling's approximation.}
\]

Using this we get:

\[
\text{MD}_\mu(x) = \frac{1}{2} \sqrt{\frac{2\alpha q}{\pi}}
\]
2. For a sequence of Bernoulli trials, let \( X \) be the number of trials required to get \( n \) successes where \( n \) is a fixed integer, if \( p \) be the probability of success for a single trial, then show that,

\[
E \left[ \frac{X}{n} \right] = \sum_{i=1}^{n-1} (-1)^{i-1} \frac{n}{n-1} \left( \frac{p}{q} \right)^i + (-\frac{p}{q})^n \ln \frac{p}{q}.
\]

**Proof:**

\[
f(x) = \binom{n-1}{x-1} p^x q^{n-x}
\]

Now,

\[
E \left( \frac{X}{n} \right) = \sum_{x=n}^\infty \frac{x}{n} \binom{n-1}{x-1} p^x q^{n-x}
\]

\[
= \left( \frac{p}{q} \right)^n \sum_{x=n}^\infty \frac{x}{n} \frac{a^x}{x}
\]

\[
= \left( \frac{p}{q} \right)^n \sum_{x=n}^\infty \frac{x-1}{x} \int_a^x t^{x-1} dt
\]

\[
= \left( \frac{p}{q} \right)^n \sum_{x=n}^\infty \frac{x-1}{x} \int_a^x t^{x-1} t^{a-x} dt
\]

\[
= \left( \frac{p}{q} \right)^n \int_a^x \sum_{x'=0}^{x-1} \frac{x}{x'} t^{x'-n} dt
\]

Let \((a-x)^{x'} = t\)

Let \( I(n-1, -n) = \int_a^x (1-t)^{n-1} dt \)

\[
= \left[ \frac{-(1-t)^{n+1}}{1-n} \right]_a^x - \int_a^x \frac{(n-1)t^{n-2}(1-t)^{n-1}}{(n-1)} dt
\]

\[
= \frac{(1-q)^{(x+1)} - 1}{(n-1)(x-1)} I[(n-2), -(n-1)]
\]

\[
I[(n-2), -(n-1)] = \frac{(a_q)^{n-2}}{n-2} I[(n-3), -(n-2)]
\]

\[
I[(n-1), -(n-1)] = \frac{(a_q)^{n-1}}{n-1} I[(n-2), -(n-2)]
\]

\[
I[(n-1), -(n-1)] = \frac{1}{n-1} \frac{(a_q)^{n-1}}{n-1} - \frac{1}{n-2} \frac{(a_q)^{n-2}}{n-2} + \frac{n-2}{n-3} \frac{(a_q)^{n-3}}{n-3}
\]

\[
I(0, -1) = \int_0^1 (1-t)^{-1} dt
\]

\[
= [-\ln(1-t)]_0^1 = -\ln \frac{p}{q}
\]
\[ I(n-1, n) = \sum_{i=1}^{n-1} \left( \frac{p}{q} \right)^{i} \frac{(-1)^{i-1}}{n-1} + (-1)^{n} \ln p \]  \hspace{1cm} (28) \\

\[ : E \left[ \frac{X}{X} \right] = \sum_{i=1}^{n-1} (-1)^{i-1} \frac{b}{n-1} \left( \frac{p}{q} \right)^{i} + \left( \frac{p}{q} \right)^{n} \ln p \quad [\text{Proved}] \]

3. Obtain the probability that in a sequence of Bernoulli trials with \( p \) for \( \alpha \) successes will occur before \( b \) failures. Hence obtain the identity

\[ \sum_{k=0}^{b-1} \left( \frac{a+b-1}{a+k} \right) p^{k} q^{b-k-1} = \sum_{k=0}^{b-1} \left( \frac{a+k-1}{k} \right) q^{k} \]

**Sol.** Let \( X \) denotes the number of failures preceding the \( \alpha \) th success, 

\[ X \sim NB(a, p) \]

\[ \therefore \text{Required probability :} \quad p \left[ X \leq b-1 \right] = \sum_{\alpha=0}^{b-1} \left( \frac{a+\alpha-1}{\alpha} \right) p^{\alpha} q^{b-\alpha} \]

Let \( \gamma \) denotes the number of failures in \( a+b-1 \) trials, 

\[ \gamma \sim Bin(a+b-1, q) \]

\[ \therefore \text{Required probability :} \quad p \left[ X \leq b-1 \right] = \sum_{\alpha=0}^{b-1} \left( \frac{a+b-1}{\alpha} \right) q^{\alpha} p^{b-\alpha} \]

\[ \therefore \sum_{\alpha=0}^{b-1} \left( \frac{a+\alpha-1}{\alpha} \right) p^{\alpha} q^{b-\alpha} = \sum_{\alpha=0}^{b-1} \left( \frac{a+b-1}{\alpha} \right) q^{\alpha} p^{b-\alpha-1} \]

\[ \therefore \sum_{\alpha=0}^{b-1} \left( \frac{a+\alpha-1}{\alpha} \right) p^{\alpha} q^{b-\alpha} = \sum_{j=0}^{b-1} \left( \frac{a+b-1}{j} \right) q^{j} p^{b-1-j} \]

\[ \therefore \sum_{k=0}^{b-1} \left( \frac{a+k-1}{k} \right) q^{k} = \sum_{k=0}^{b-1} \left( \frac{a+b-1}{k} \right) q^{b-1-k} p^{k} \quad [\text{Proved}] \]
Mode of a Negative Binomial Distribution:

Let \( X \sim NB(n, p) \), the p.m.f. of \( X \) is \( f(x) = \binom{x+n-1}{x} p^n q^x \).

The mode of the distribution is the value (s) of \( x \) for which \( f(x) \) is maximum. For \( x = 1, 2, \ldots, \infty \)

Let us consider the ratio,

\[
\frac{f(x)}{f(x-1)} = \frac{(x+n-1) p^n q^x}{(x+n-2) p^n q^{x-1}} = \frac{n+x-1}{x}, q
\]

\[
\therefore \frac{f(x)}{f(x-1)} \geq 1 \text{ according as } (n+x-1)(1-p) \geq x
\]

\[\Rightarrow (n-1)(1-p) \geq x \]

\[\Rightarrow x \leq (n-1) \frac{(1-p)}{p}
\]

\[
\therefore f(x) \geq f(x-1) \text{ according as } x \leq (n-1) \frac{(1-p)}{p}.
\]

Case I:

\( (n-1) \frac{(1-p)}{p} \) is an integer = \( k \) (say)

Now, \( x = k \)

\[\Rightarrow f(k) = f(k-1) \quad \ldots \ldots \quad 1 \]

For, \( x < k \)

\[\Rightarrow f(x) > f(x-1) \quad \text{for} \quad x = 1, 2, \ldots, k-1. \]

\[\Rightarrow f(1) < f(2) < \ldots \ldots < f(k-1) \quad \ldots \ldots \quad 2\]

For, \( x > k \)

\[\Rightarrow f(x) < f(x-1) \quad \text{for} \quad x = k+1, k+2, \ldots, \infty. \]

\[\Rightarrow f(k) > f(k+1) > \ldots \ldots > \infty \quad \ldots \ldots \quad 3\]

Combining 1, 2 & 3, we get:

\[f(1) < f(2) < \ldots \ldots < f(k-1) = f(k) > f(k+1) > \ldots \ldots .\]

\[
\therefore f(x) \text{ is maximum at } \alpha = k \text{ and } \alpha = k-1.
\]

\[\therefore \text{The mode of the distribution are at }\]

\((n-1) \frac{(1-p)}{p} \) and \((n-1) \frac{(1-p)}{p} - 1\).
Case II: If \( \frac{(r-1)(1-p)}{p} \) is not an integer.

Let, \( k_0 = \left[ \frac{(r-1)(1-p)}{p} \right] \) = largest integer obtained in \( \frac{(r-1)(1-p)}{p} \)

Obviously,
\[
\frac{\alpha}{p} < \frac{(r-1)(1-p)}{p} < \frac{\alpha + 1}{p}
\]

Now, \( \alpha = k_0 \),
\[
\Rightarrow \alpha < \frac{(r-1)(1-p)}{p}
\]
\[
\Rightarrow f(k_0) > f(k_0 - 1) \quad \text{for } k_0 = 1, 2, \ldots, k_0 - 1
\]
\[
\Rightarrow f(0) < f(1) < \ldots < f(k_0 - 1) \quad \text{for } k_0 = 1, 2, \ldots, k_0 - 1
\]

\[
\alpha > k_0
\]
\[
\Rightarrow \alpha > k_0 + 1 > \frac{(r-1)(1-p)}{p}
\]
\[
\Rightarrow f(\alpha) < f(\alpha - 1) \quad \text{for } \alpha = k_0 + 1, k_0 + 2, \ldots, \infty
\]
\[
\Rightarrow f(k_0) > f(k_0 + 1) > \ldots > \infty
\]

Combining 4, 5, and 6, we get
\[
f(0) < f(1) < \ldots < f(k_0 - 1) < f(k_0) > f(k_0 + 1) > \ldots
\]

\( f(\alpha) \) is maximum at \( \alpha = k_0 \).

**Factorial moment generating function:**

Let \( X \sim \text{NB}(r, p) \)

PGF of \( X \) is given by,
\[
P_X(t) = E(t^X), \text{ since } |t| < 1,
\]
\[
= \sum_{x=0}^{\infty} (2 + r - 1) p^x q^x t^x
\]

\[
= p^r (1 - qt)^{-r} \quad \text{[ : } |t| < \frac{1}{q} \]

Now, factorial moment generating function is
\[
P_X(1 + t) = p^r (1 - q - qt)^{-r}
\]
\[
= (1 - \frac{q}{p} t)^{-r} = \sum_{x=0}^{\infty} \left( \frac{x + r - 1}{\alpha} \right) \left( \frac{qt}{p} \right)^x
\]
\[
= \sum_{x=0}^{\infty} \left( \frac{x + r - 1}{\alpha} \right) \left( \frac{q}{p} \right)^x \frac{x!}{\alpha!} t^x
\]
\[ M[2] = \text{Coefficient of } \frac{t^2}{2!} \text{ in the expansion of } f(x) \]

\[ = (a+n-1)x \left( \frac{a}{p} \right)^2 \]

(\text{Factorial moment} \rightarrow \text{Raw moment} \rightarrow \text{Central moment})

\[ M_1' = E(X) = \frac{npq}{p} = \text{mean} \]
\[ = n(n+1) \frac{aq}{p^2} + \frac{na}{p} \]
\[ = n(n+1)(n+2) \frac{aq^2}{p^3} + 3n(n+1) \frac{aq}{p^2} + \frac{na}{p} \]
\[ = n(n+1)(n+2)(n+3) \left( \frac{aq}{p} \right)^3 + 6n(n+1)(n+2) \left( \frac{aq}{p} \right)^2 \]
\[ + 7n(n+1) \left( \frac{aq}{p} \right) + \frac{na}{p} \]

\( \star 4. \) Suppose the probability that a workman chosen at random will possess accident proneness of intensity \( \lambda \) to \( \lambda + d\lambda \) and that he produces \( x \) accidents is

\[ f(x) d\lambda p(x, \lambda), \]

where
\[ f(x) = \frac{\lambda^x}{\Gamma(x)} \exp[-\lambda] \lambda^{x-1}, \quad 0 < \lambda < \infty, \]

and
\[ p(x, \lambda) = \frac{\exp[\lambda\lambda^x]}{x!}, \quad x = 0, 1, 2, \ldots \]

Show that the probability that a workman chosen at random (with unknown accident proneness) will produce \( x \) accidents follows a negative Binomial distribution, where \( p = \frac{\lambda}{(\lambda+1)} \) and \( n = \alpha \).

Solution:

\[ f(x) = \lim_{d\lambda \to 0} \frac{F(\lambda+d\lambda)-F(\lambda)}{d\lambda} \]

\[ = \lim_{d\lambda \to 0} \frac{p(\lambda < X < \lambda+d\lambda)}{d\lambda} \]

For small \( d\lambda \), (20)

\[ p [ \lambda < X < \lambda + d\lambda ] \approx f(x) d\lambda \text{ is the probability element.} \]
The p.m.f. of $X$ is,
\[ f(x) = \int_0^\infty p(x, \lambda) f(\lambda) d\lambda \]
\[ = \frac{\gamma^\alpha}{\lambda^\alpha \Gamma(\alpha)} \int_0^\infty e^{-(1+\gamma)\lambda} \lambda^\alpha e^{-\lambda} \lambda^{-\alpha} d\lambda \]
\[ = \frac{\gamma^\alpha}{\lambda^\alpha \Gamma(\alpha)} \int_0^\infty e^{-(1+\gamma)\lambda} \frac{\lambda^\alpha}{(1+\gamma)^{\alpha+1}} d\lambda \]
\[ = \frac{\gamma^\alpha}{\lambda^\alpha \Gamma(\alpha)} \Gamma(\alpha+\alpha) \frac{1}{(1+\gamma)^{\alpha+1}}, \quad \alpha = 0, 1, 2, \ldots \]
\[ = \frac{(\alpha+\alpha-1)!}{\alpha! (\alpha-1)!} \frac{\gamma^\alpha}{(1+\gamma)^{\alpha+1}} \frac{1}{(1+\gamma)^{\alpha-1}} \]
\[ = \binom{\alpha+\alpha-1}{\alpha} \left( \frac{\gamma}{1+\gamma} \right)^\alpha \left( \frac{1}{1+\gamma} \right)^\alpha \]

Here $X \sim NB(n, p)$, where $n = \alpha$, $p = \left( \frac{\gamma}{1+\gamma} \right)$.

**Remark:** In Bin($n$, $p$) distribution, we count the no. of successes in $n$ independent Bernoulli trials whereas in NB($n$, $p$), we count the no. of trials required to get the $n$th success. Therefore, Negative Binomial is also inverse Binomial Sampling.
A discrete random variable $X$ with mass points $0, 1, 2, \ldots$ is said to follow the Poisson distribution with parameter $\lambda > 0$ if its p.m.f. is given by:

$$f(x) = \left\{ \begin{array}{ll}
\frac{e^{-\lambda} \lambda^x}{x!}, & x = 0, 1, 2, \ldots \\
0 & \text{otherwise}
\end{array} \right.$$ 

We denote $X \sim P(\lambda)$.

Example: Some real-life situations where the Poisson law is appropriate:

i) Distribution of number of misprints on a certain page of an encyclopedia.

ii) No. of calls connection during an hour in a busy telephone exchange.

iii) No. of total road accidents during hour in a five point junction.

iv) No. of defects in a piece of cloth of specified length.

v) No. of defective items in a factory.

Uses: Poisson Distribution is used to model the probability of occurrence of rare events.

Remark: The number of misprints on a page of a book is the probability that each letter, typed on a page will be misprinted is small and the no. of letters on a page is quite large. But, in the first proof of a book, the probability that a letter typed is misprinted is not small enough and consequently "the no. of misprints on a page of a book" will not follow approximately a Poisson distribution.

In practice, great care has to be taken to avoid erroneously application to Poisson distribution for counting variables. For e.g., in studying the distribution of deaths of insects form a cluster which is inconsistent with the assumption of Poisson Process.
Poisson Process: Suppose that we are observing the occurrence of an event in time, a quantity \( M(t) \) such that:

i) The probability that exactly one happening will occur in a small-time interval of length \( t \) is \( \lambda t \).

ii) The probability that more than one happening in a small-time interval of length \( t \) is negligible.

iii) The no. of happenings in non-overlapping intervals are independent.

Then the no. of occurrences in the interval of the length \( T \) follows a “Poisson distribution” with mean \( \lambda T \). Here, \( \lambda \) is known as the mean rate of occurrence.

Then the no. of occurrences in the time interval of length \( \frac{T}{2} \) follows “Poisson Distribution” with mean \( \frac{\lambda T}{2} \).

From the conditions of Poisson process, certain random experiments revolving around occurrence of an event in time, space or length can be realistically modelled by Poisson Distribution.

Result: - Let \( X \sim \text{Bin}(n, p) \), suppose the following conditions are satisfied:

i) \( n \to \infty \) (i.e. the no. of trials is very large).

ii) \( p \to 0 \) (i.e. the probability of success is very small).

iii) \( np = \lambda \) (finite) [i.e. the average number of trials is finite].

Then the distribution of \( X \) will converge to a Poisson distribution with parameter \( \lambda \).

Proof: For some fixed small \( \alpha \),

\[
P[X = \alpha] = \binom{n}{\alpha} p^\alpha (1-p)^{n-\alpha}, \quad \alpha = 0(1)n,
\]

\[
= \frac{n(n-1) \cdots (n-\alpha+1)}{\alpha!} \left( \frac{\lambda}{n} \right)^{\alpha} \left( 1 - \frac{\lambda}{n} \right)^{n-\alpha}
\]

\[
= \frac{1 - (1 - \frac{\lambda}{n})^{\alpha}}{\alpha!} \left( 1 - \frac{\lambda}{n} \right)^{n-\alpha}
\]

As \( n \to \infty \), \( p \to 0 \), \( np = \lambda \),

\[
(1 - \frac{\lambda}{n})^{\alpha} \to 1, \quad \text{and} \quad (1 - \frac{\lambda}{n})^n \to e^{-\lambda}
\]

\[
\Rightarrow \lim_{n \to \infty} P[X = \alpha] = \frac{e^{-\lambda} \lambda^\alpha}{\alpha!}
\]

\[
\text{limiting case of Bin. distn when } n \to \infty, p \to 0 \text{ but } np = \lambda \text{ (finite)}
\]

\[
\text{can be approximated by Pois}(\lambda).
\]
Another method: \( x \sim \text{Bin}(n, p) \)

\[
M_x(t) = (q + pe^t)^n
= \left\{ \begin{array}{l}
1 - p(1-e^t)^n \\
1 - \frac{np}{n}(1-e^t)^n \\
1 - \frac{2}{n}(1-e^t)^n
\end{array} \right.
\]

\[
\therefore \lim_{n\to\infty} M_x(t) = \exp \left[ -\lambda (1-e^t) \right]
= e^{\lambda (e^t-1)}
\]

which is the mgf of a Poisson distribution parameter \( \lambda \),
hence, by uniqueness of mgf, as \( n \to \infty, p \to 0, np = \lambda \),
\( x \) is a Poisson R.V.

\[\square \text{Result:} - \text{Let } x \sim \text{NB}(n, p), \text{ suppose the following conditions are satisfied,} \]

i) \( n \to \infty \)

ii) \( p \to 0 \)

iii) \( np = \lambda \) (finite)

Then the distn. of \( x \) will converge to a Poisson distn. with parameter \( \lambda \).

\[\text{Proof:} - \text{for some fixed } \alpha, \]

\[
P[X = \alpha] = \binom{\alpha + n - 1}{\alpha} p^\alpha q^{n-\alpha}, \quad \alpha = 0, 1, 2, \ldots
\]

\[
\begin{align*}
&= \frac{(\alpha+n-1) \cdot \ldots \cdot \alpha}{\alpha!} \cdot p^\alpha q^{n-\alpha} \\
&= \frac{(\frac{\alpha-1}{n} + 1) \cdot \ldots \cdot 1}{\alpha!} \cdot p^\alpha \cdot q^{n-\alpha}
\end{align*}
\]

\[
\therefore \text{as } n \to \infty, \quad (\frac{\alpha-1}{n} + 1) \cdot \ldots \cdot 1 \to 1
\]

\[
\text{& } (1 - \frac{p}{n})^n \to e^{-\lambda}
\]

\[
\therefore f(x) = e^{-\lambda} \frac{\lambda^x}{x!}
\]
Another method: Let \( X \sim \text{NB}(n, p) \)

Suppose, \( n \rightarrow \infty \), \( q \downarrow 0 \) and \( q = 1 - p \) \( \Rightarrow nq = \lambda \) (finite)

P.G.F. of \( X \) is

\[
P_X(t) = p^n (1-qt)^{-n}
\]

\[
= (1-q)^n (1-\frac{at}{p})^{-n} = (1-\frac{2}{p})^n (1-\frac{a}{p})^{-n}
\]

\[
\Rightarrow e^{-\lambda} \cdot e^{\lambda t} = e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!}
\]

\[P[X=x] = \frac{e^{-\lambda} \cdot \lambda^x}{x!}
\]

[ Poisson Distn. as a limiting case of negative Binomial Distribution ]

Recurrence Relation Regarding Probability:

\[
\frac{P(X = x)}{P(X = x-1)} = \frac{e^{-\lambda} \cdot \frac{\lambda^x}{x!}}{e^{-\lambda} \cdot \frac{\lambda^{x-1}}{(x-1)!}} = \frac{\lambda}{x}
\]

\[
P(X = x) = \frac{\lambda^x}{x!} \cdot P(X = x-1), ~x = 1, 2, 3, \ldots
\]

Factorial Moments:

Let \( X \sim \text{Pois}(\lambda) \), then the \( r \)th order factorional moment of \( X \) is given by

\[
\mu_r = E[(X)^r] = E[X(X-1) \cdots (X-r+1)]
\]

\[
= \sum_{x=0}^{\infty} (x)^r \cdot \frac{e^{-\lambda} \cdot \lambda^x}{x!}
\]

\[
= \sum_{x=0}^{\infty} (x)^r \cdot e^{-\lambda} \cdot \frac{\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^x}{x!} \cdot \sum_{x=r}^{\infty} \frac{\lambda^{x-r}}{(x-r)!}
\]

\[
= e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} \cdot \frac{\lambda^{x-r}}{(x-r)!}
\]

\[
= e^{-\lambda} \cdot \sum_{x=0}^{\infty} \frac{\lambda^{x}}{(x!)(x-r)!} = \mu_r
\]

Note, we can get raw moments from factorial moments and central moments from raw moments.
\[ M_1 = \mu_{[1]} = \lambda \]
\[ M_2 = \mu_{[2]} + \mu_{[1]} = \lambda^2 + \lambda \]
\[ M_3 = \mu_{[3]} + 3\mu_{[2]} + \mu_{[1]} = \lambda^3 + 3\lambda^2 + \lambda \]
\[ M_4 = \mu_{[4]} + 6\mu_{[3]} + 7\mu_{[2]} + 3\mu_{[1]} = \lambda^4 + 6\lambda^3 + 7\lambda^2 + 3\lambda + 1 \]

\[ \therefore E(X) = M_{[1]} = \lambda \]
\[ \therefore \text{Var}(X) = M_{[2]} + \mu_{[1]} - [M_{[1]}]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \]

Central moments:

\[ M_{[1]} = M_1 = \mu_1 = \lambda \]
\[ M_2 = \mu_2 = \lambda^2 \]
\[ M_3 = \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3 = \lambda^3 + 3\lambda^2 + \lambda - 3(\lambda^2 + \lambda) + 2\lambda^3 = \lambda^3 \]
\[ M_4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 3\mu_1^4 = 6\lambda^5 + \lambda^4 \]

Skewness:

\[ \beta_1 = \frac{M_3}{M_2^{3/2}} = \frac{\lambda^3}{\lambda^{3/2}} = \frac{\lambda^{3/2}}{\lambda} = \frac{\lambda}{\lambda^{3/2}} = \frac{1}{\lambda^{3/2}} \]

\[ \therefore \gamma_1 = \frac{1}{\sqrt{\lambda}} \], since \( \lambda > 0 \), Thus the Poisson distribution is positively skewed.

Kurtosis:

\[ \beta_2 = \frac{M_4}{M_2^2} = \frac{3\lambda^4 + \lambda^2}{\lambda^4} = 3 + \frac{1}{\lambda} \]

\[ \therefore \gamma_2 = \beta_2 - 3 = 3 + \frac{1}{\lambda} - 3 = \frac{1}{\lambda} > 0 \]

Thus the distribution is leptokurtic.
Probability Generating Function: \[ P_G(t) = E(t^X) = \sum_{x=0}^{\infty} t^x \frac{e^{-\lambda} \lambda^x}{x!} = e^{\lambda(t-1)} \]

Factorial moment Generating function: \[ P(1+t) = e^{\lambda t} = \sum_{x=0}^{\infty} \frac{(\lambda t)^x}{x!} \]

Coefficient of \[ t^{x-1} / x! = \lambda^x = \lambda^x \]

Moment Generating Function: \[ M_G(t) = M(t) = e^{\lambda(e^t-1)} \quad \forall \ t \in \mathbb{R} \]

\[ = e^{\lambda \left( \sum_{x=1}^{\infty} \frac{t^x}{x!} \right)} \]

\[ = 1 + \lambda \frac{\lambda t^2}{2} + \frac{\lambda^2 t^3}{3} + \frac{\lambda^3 t^4}{4} + \cdots + \frac{\lambda^x}{x!} \]

\[ = 1 + \lambda \left( \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \cdots + \frac{t^x}{x!} \right) \]

\[ + \frac{\lambda^3}{3!} \left( \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \cdots + \frac{t^x}{x!} \right) + \cdots + \frac{\lambda^x}{x!} \]

\[ = e^{\lambda \left( \frac{t^2}{2} + \frac{t^3}{3} + \frac{t^4}{4} + \cdots + \frac{t^x}{x!} \right)} \]

\[ \therefore \mu_1' = \lambda ; \quad \mu_2' = \lambda + \lambda^2 ; \quad \mu_3' = \lambda + 3\lambda^2 + \lambda^3 ; \quad \mu_4' = \lambda^4 + 7\lambda^3 + 6\lambda^2 + \lambda \]

Central moment Generating Function: \[ M_x(t) = E \left[ e^{t(x-\mu)} \right] \]

\[ = e^{-\lambda t} E(e^{t(x-\mu)}) \]

\[ = e^{-\lambda t} \cdot e^{\lambda(e^t-1)} \]

\[ = e^{\lambda(e^t-1)-t} = e^{\lambda \left( \frac{t^2}{2} + \frac{t^3}{3} + \cdots \right)} \]

\[ = 1 + \lambda \left( \frac{t^2}{2} + \frac{t^3}{3} + \cdots \right) + \frac{\lambda^2}{2} \left( \frac{t^2}{2} + \frac{t^3}{3} + \cdots \right) + \cdots \]

\[ \therefore \mu_1 = \lambda ; \quad \mu_2 = \lambda^2 \lambda^3 = \lambda \lambda^3 , \quad \mu_4 = \lambda + 3\lambda^2 \]
Cumulant Generating Function: 

\[ k(t) = CGF = \ln M(t) = \ln \lambda(e^t - 1) = \lambda(e^t - 1) = \sum_{n=1}^{\infty} \frac{\lambda^i \cdot t^n}{i!} \]

\[ \Rightarrow k_i = \lambda \quad \forall \quad i = 1, 2, \ldots, \infty. \]

All the cumulants of Poisson distribution are equal.

- \[ \mu_1 = k_1 = \lambda \]
- \[ \mu_2 = k_2 = \lambda \]
- \[ \mu_3 = k_3 = \lambda \]
- \[ \mu_4 = k_4 + 3k_2^2 = \lambda + 3\lambda^2 \]

Recursion Relation for Central moments:

\[ \mu_0 = E(x - \mu)^0 = \sum_{x=0}^{\infty} (x - \mu)^0 \cdot e^{-\lambda} \cdot \frac{\lambda^x}{x!} \]

Now, differentiating both sides w.r.t. \( \lambda \), we get,

\[ \frac{d\mu_0}{d\lambda} = \sum_{x=0}^{\infty} (-x) (x - \mu)^{n-1} \cdot e^{-\lambda} \cdot \frac{\lambda^n}{x!} \]

\[ + \sum_{x=0}^{\infty} (x - \mu)^n \cdot \frac{1}{x!} \left[ -e^{-\lambda} \cdot \frac{\lambda^n}{x!} + \frac{\lambda}{x} \sum_{x=0}^{\infty} (x - \mu)^n \cdot e^{-\lambda} \cdot \frac{\lambda^n}{x!} \right] \]

\[ = -\lambda \sum_{x=0}^{\infty} (x - \mu)^{n-1} \cdot e^{-\lambda} \cdot \frac{\lambda^n}{x!} + \frac{1}{\lambda} \sum_{x=0}^{\infty} (x - \mu)^n \cdot e^{-\lambda} \cdot \frac{\lambda^n}{x!} \]

\[ = -\lambda \mu_{n-1} + \frac{1}{\lambda} \mu_n \]

\[ \therefore \mu_{n+1} = \lambda \left[ \frac{d}{d\lambda} \mu_n + \mu_1 \mu_{n-1} \right], \quad n = 1, 2, \ldots. \]

- Putting \( n = 1 \), \[ \mu_2 = \lambda \left[ 0 + 1 \right] = \lambda \]
- Putting \( n = 2 \), \[ \mu_3 = \lambda \left[ 1 + 2 \times 0 \right] = \lambda \]
- Putting \( n = 3 \), \[ \mu_4 = \lambda \left[ 1 + 3 \lambda \right] \]
  \[ = \lambda + 3\lambda^2. \]
Mode of the Poisson Distribution:

\[ x \sim \mathcal{P}(\lambda) \]

\[ f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ \therefore \frac{f(x)}{f(x-1)} = \frac{x}{\lambda}, \quad x = 1, 2, \ldots, \infty \]

\[ \therefore \frac{f(x)}{f(x-1)} \leq 1 \text{ according as } x \leq \lambda. \]

**Case I:** \( \lambda \) is an integer.

\[ f(x-1) < f(x) \text{ for } x = 1, 2, \ldots, \lambda - 1 \]

\[ f(x-1) = f(x) \text{ for } x = \lambda \]

\[ f(x-1) > f(x) \text{ for } x = \lambda + 1, \lambda + 2, \ldots \]

\[ \therefore f(0) < f(1) < f(2) < \ldots < f(\lambda-1) = f(\lambda) > f(\lambda+1) > f(\lambda+2) > \ldots \]

Thus, \( \lambda \) and \( \lambda - 1 \) are the two modes of the distribution.

**Case II:** \( \lambda \) is not an integer, \( [\lambda] = m \).

\[ f(x-1) < f(x) \text{ for } x = 1, 2, \ldots, m \]

\[ f(x-1) > f(x) \text{ for } x = m+1, m+2, \ldots \]

\[ \therefore m \text{ is the mode of the distribution, where } \lambda \text{ is not an integer, i.e., the largest integer cannot exceed } \lambda. \]

**Mean Deviation about Mean:**

\[ x \sim \mathcal{P}(\lambda), \quad \mu = \lambda \]

\[ \therefore M.D. \mu(x) = E[|x - \mu|] \]

\[ = \sum_{x=0}^{\infty} |x - \mu| \cdot \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ = \sum_{x=0}^{\infty} (x - \mu) \cdot \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ = \sum_{x=0}^{\infty} (x - \lambda) \cdot \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ = \sum_{x=0}^{\lambda} (x - \lambda) \cdot \frac{e^{-\lambda} \lambda^x}{x!} + \sum_{x=\lambda+1}^{\infty} (x - \lambda) \cdot \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ = \lambda \cdot \sum_{x=0}^{\lambda} \frac{e^{-\lambda} \lambda^x}{x!} - \sum_{x=\lambda+1}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \]

\[ \text{[let } x-\lambda = \xi \text{ in R.H.S]} \]

\[ = \lambda \cdot [ \sum_{x=0}^{\lambda} \frac{e^{-\lambda} \lambda^x}{x!} - \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} ] \]

\[ = \lambda \cdot [ \frac{e^{-\lambda} \lambda^\lambda}{\lambda!} - \frac{e^{-\lambda} \lambda^{\infty}}{\lambda!} ] \]

\[ = \lambda \cdot \frac{e^{-\lambda} \lambda^\lambda}{\lambda!} - \frac{e^{-\lambda} \lambda^{\infty}}{\lambda!} \]

\[ \text{for large } \lambda, \quad \alpha_0 \to \lambda, \quad \text{and } M.D. \mu(x) \to \frac{2e^{-\lambda} \lambda^{\alpha_0+1}}{\lambda!} \]
Using Stirling's approximation for large $n$,
\[
M_{\mu}(X) = \frac{2 \cdot e^{-\lambda} \cdot \lambda^{n+1}}{\sqrt{2\pi} \cdot e^{-\lambda} \cdot \lambda + \frac{1}{2}}
\]
\[
= \frac{2\lambda}{\sqrt{\pi}}
\]
Hence, standard deviation $= \sqrt{\lambda}$

\[
\therefore \frac{M_{\mu}(X)}{s.d.} = \sqrt{\frac{2\lambda}{\pi}} \cdot \frac{1}{\sqrt{\lambda}} = \sqrt{\frac{2}{\pi}}
\]

\[\star \textbf{Problem 1.} \quad \text{If } X \sim P(\lambda = 1), \text{ then show that } M_{\mu}(X) \text{ is equal to } \left( \frac{2}{e} \times \text{s.d.} \right). \]

\[\text{Soln.} \implies M_{\mu}(X) = E[X-1], \quad \lambda = 1
\]
\[
= E\left[\frac{X-1}{2}\right]
\]
\[
= 2 \sum_{x=0}^{\infty} (1-x) \cdot f(x)
\]
\[
= 2 \cdot f(0)
\]
\[
= 2 \cdot e^{-1} \cdot 1^0 \cdot 0!
\]
\[
= \frac{2}{e} \cdot \sqrt{\lambda} \quad \text{[as } \sqrt{\lambda} = 1]\n\]
\[
= \frac{2}{e} \times \text{s.d.} \quad \text{(Proved)}
\]

\[\star \textbf{Problem 2.} \quad \text{If } X \sim P(\lambda), \text{ Find } E\left[\frac{1}{1+x}\right]. \]

\[\text{Soln.} \quad E\left(\frac{1}{1+x}\right) = \sum_{j=0}^{\infty} \frac{1}{(1+j)} \cdot e^{-\lambda} \cdot \frac{\lambda^j}{j!}
\]
\[
= \frac{1}{\lambda} \sum_{j=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{j+1}}{(j+1)!}
\]
\[
= \frac{1}{\lambda} \left[ \sum_{j=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^{j}}{j!} - e^{-\lambda} \right] \quad ; \quad \tilde{j} = j + 1
\]
\[
= \frac{1}{\lambda} \left[ 1 - e^{-\lambda} \right]
\]
Problem 3. If \( X \sim P(\lambda) \) then show that

\[
E[Y] = \lambda E[X]
\]

\[
E[X^2] = \sum_{j=1}^{\infty} j^2 \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \lambda \sum_{j=1}^{\infty} j^2 \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \lambda \sum_{j=0}^{\infty} (j+1) \frac{e^{-\lambda} \lambda^{j+1}}{(j+1)!}
\]

\[
= \lambda \sum_{j=0}^{\infty} j \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \lambda E[X] = \lambda \lambda = \lambda^2
\]

Problem 4. If \( X \sim P(\lambda) \) then show that \( E(X^n) = \lambda E(X+1)^{n-1} \).

Hence find \( E(X) \) & \( Var(X) \).

\[
E(X^n) = \sum_{j=1}^{\infty} j^n \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \sum_{j=1}^{\infty} j^n \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \lambda \sum_{j=1}^{\infty} (j-1)^n \frac{e^{-\lambda} \lambda^j}{(j-1)!}
\]

\[
= \lambda \sum_{j=0}^{\infty} j^n \frac{e^{-\lambda} \lambda^j}{j!}
\]

\[
= \lambda \lambda^{n-1} \sum_{j=0}^{\infty} \frac{(j+1)^{n-1}}{j!} e^{-\lambda} \lambda^j
\]

\[
= \lambda \lambda^{n-1} E(X+1)^{n-1}
\]

Now putting, \( n=1 \),

\[
E(X) = \lambda
\]

\[
E(X^2) = \lambda E(X+1)
\]

\[
= \lambda E(X) + \lambda
\]

\[
= \lambda^2 + \lambda
\]

\[
\therefore Var(X) = E(X^2) - E(X)^2
\]

\[
= \lambda^2 + \lambda - \lambda^2
\]

\[
= \lambda
\]
Problem 5. If \( X \sim \mathcal{P}(\lambda) \) then show that:

\[
\mu_{n+1} = \mu_0 \left[ \binom{n}{0} \mu_{n-1} + \binom{n}{1} \mu_{n-2} + \cdots + \binom{n}{n} \mu_0 \right]
\]

**Solution:**

\[
\mu_{n+1} = E(X-\mu)^{n+1} = \sum_{j=0}^{\infty} (j-\mu)^{n+1} f(j)
\]

\[
= \sum_{j=0}^{\infty} (j-\mu)^{n+1} \frac{\lambda^j}{j!}
\]

\[
= \sum_{j=0}^{\infty} (j-\mu)^{n+1} \frac{e^{-\lambda} \lambda^j}{j!} - \lambda \sum_{j=0}^{\infty} (j-\mu)^{n} f(j)
\]

\[
= \lambda \sum_{j'=-\infty}^{\infty} \left[ \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} (j'-\mu)^n f(j') \right] - \lambda \mu^n
\]

\[
= \lambda \sum_{i=0}^{\infty} \left( \binom{i}{0} \right) \sum_{j'=0}^{\infty} (j'-\mu)^n f(j') - \lambda \mu^n
\]

\[
= \lambda \sum_{i=0}^{\infty} \left( \binom{i}{0} \right) \mu_0 - \lambda \mu^n
\]

\[
= \lambda \sum_{i=0}^{\infty} \left( \binom{i}{0} \right) \mu_{n-1} - \lambda \mu^n
\]

**Problem 6.** For the Poisson distribution with parameter \( \lambda \), S.T.

\[ E(k, X, e^{-kX}) = \lambda k e \lambda (e^{-k-1}) - k \quad k \text{ being a real constant} \]

**Solution:**

\[ E(k, X, e^{-kX}) = \sum_{\alpha=0}^{\infty} e^{-\lambda} \frac{\lambda^\alpha}{\alpha!} \frac{\lambda^x e^{-\lambda}}{\alpha!} \]

\[ = k \lambda e^{-\lambda} \sum_{\alpha=0}^{\infty} \frac{e^{-k\lambda}}{(\alpha-1)!} \]

\[ = e^{-\lambda} k \lambda e^{-k} \lambda \lambda \lambda k \left[ \sum_{\alpha=1}^{\infty} \frac{(e^{-k-1})^{\alpha-1}}{(\alpha-1)!} \right]
\]

\[ = e^{-\lambda} k \lambda e^{-k} \lambda \lambda k \left[ \sum_{\alpha'=0}^{\infty} \frac{(e^{-k-1})^{\alpha'}}{\alpha'!} \right] \]

\[ = \lambda k \lambda e^{-k} \lambda (e^{-k-1}) - k \quad (\text{Proved}) \]

\( k \) being a real constant.
Problem 7. If $x_1$ and $x_2$ be two independent Poisson random variables with common expected value $\lambda$, then show that
the probability that $x_1x_2$ is even is $\frac{1}{4} \left( 3 + 2e^{2\lambda} - e^{4\lambda} \right)$.

Solution \( x_1 \sim P(\lambda), \quad x_2 \sim P(\lambda) \).

\[
P(x_1x_2 = \text{even}) = 1 - P(x_1x_2 = \text{odd})
\]
\[
= 1 - \left[ P(x_1 = \text{odd}) \times P(x_2 = \text{odd}) \right]
\]
\[
\text{as} \quad x_1, x_2 \text{ are ind.}
\]
\[
P(x_1 = \text{odd}) = \frac{\lambda^3 e^{-\lambda}}{3!} + \frac{\lambda^5 e^{-\lambda}}{5!} + \cdots \infty
\]
\[
= e^{-\lambda} \left[ \frac{\lambda^3}{3!} + \frac{\lambda^5}{5!} + \cdots \infty \right]
\]
\[
= e^{-\lambda} \left[ \frac{e^\lambda - e^{-\lambda}}{2} \right]
\]
\[
= \frac{1 - e^{-2\lambda}}{2}
\]

:. \( 0 \) gives

\[
P(x_1x_2 = \text{even}) = 1 - \left( \frac{1 - e^{-2\lambda}}{2} \right) \left( \frac{1 - e^{-2\lambda}}{2} \right)
\]
\[
= 1 - \frac{1 - e^{-2\lambda} - e^{-2\lambda} + e^{-4\lambda}}{4}
\]
\[
= \frac{1}{4} \left( 3 + 2e^{-2\lambda} - e^{-4\lambda} \right)
\]

Problem 8. If $x \sim P(\lambda)$ with mean $\mu = \lambda$ and variance $\sigma^2 = \lambda$, show that

\[
P \left[ 0 \leq x \leq 2\lambda \right] \geq 1 - \frac{1}{\lambda} \quad \text{(Using Chebyshev's inequality)}
\]

Solution \( x \sim P(\lambda) \)

\[
\mu = \lambda, \quad \sigma^2 = \lambda,
\]

From Chebyshev's inequality, we know,
\[
P \left( \left| \frac{x - \mu}{\sigma} \right| \leq t \right) > 1 - \frac{1}{t^2}, \quad \text{where} \quad \mu, \sigma < \infty.
\]

So,
\[
P \left[ 0 \leq x \leq 2\lambda \right] = P \left[ \left| x - \mu \right| \leq \lambda \right]
\]
\[
= P \left[ \left| \frac{x - \mu}{\sqrt{\lambda}} \right| \leq \sqrt{\lambda} \right]
\]
\[
= P \left[ \left| \frac{x - \mu}{\sigma} \right| \leq \sqrt{\lambda} \right] \quad \text{[\( \frac{\sigma}{\sqrt{\lambda}} = 1 \)]}
\]
\[
> 1 - \frac{1}{\left( \sqrt{\lambda} \right)^2} = 1 - \frac{1}{\lambda},
\]

Hence the proof is complete.
Problem 9. Let the p.m.f. be positive on and only on the non-negative integers $0, 1, 2, \ldots$; given that $f(x) = \frac{\lambda^x}{x!} e^{-\lambda}$ for $x = 1, 2, \ldots$.

Determine $f$.

\[ f(x) = \frac{\lambda^x}{x!} f(x-1) \quad \text{for} \quad x = 1, 2, \ldots \]

\[ f(0) = \lambda f(0) \]

\[ f(1) = \frac{\lambda^2}{2} f(1) = \frac{\lambda^2}{2} f(0) \]

\[ f(2) = \frac{\lambda^3}{3!} f(2) = \frac{\lambda^3}{3!} f(0) \]

\[ f(3) = \frac{\lambda^4}{4!} f(3) = \frac{\lambda^4}{4!} f(0) \]

\[ \vdots \]

\[ f(x) = \frac{\lambda^x}{x!} f(x) = \frac{\lambda^x}{x!} f(0) \quad \text{for} \quad x = 1, 2, \ldots \]

We know \( \sum_{x=0}^{\infty} f(x) = 1 \) \[ \Rightarrow \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} f(0) = 1 \]

\[ f(0) = e^{-\lambda} \]

\[ f(1) = \frac{\lambda e^{-\lambda}}{1!} \quad \Rightarrow \quad f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad ; \quad \lambda \sim P(\lambda).

Another way:\[ \frac{f(x)}{f(x-1)} = \frac{\lambda}{x} \quad ; \quad x = 1, 2, \ldots \]

\[ = \frac{\lambda x e^{-\lambda}}{x!} \]

\[ = \frac{\lambda x e^{-\lambda}}{(x-1)!} \]

\[ \Rightarrow \quad f(x) = \frac{\lambda x e^{-\lambda}}{(x-1)!} \quad \text{if} \quad f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for} \quad x = 1, 2, \ldots \]

If we can show \( \sum_{x=0}^{\infty} f(x) = 1 \), then we can say $f(x)$ is the required p.m.f.

\[ \sum_{x=0}^{\infty} \frac{\lambda^x e^{-\lambda}}{x!} = e^{-\lambda}, e^\lambda = 1 \quad \text{(checked)}.

So, \[ f(x) = \frac{\lambda^x e^{-\lambda}}{x!} \quad ; \quad x \sim P(\lambda). \]
Distribution function in terms of Incomplete Gamma function:

\( X \sim P(\lambda) \) then for any non-negative integer \( k \),

\[
P[X \leq k] = \frac{1}{\Gamma(k+1)} \int_0^\infty e^{-u} u^k \, du \quad \text{[Incomplete Gamma function]}
\]

**Proof:**

\[
P[X \leq k] = \sum_{\alpha = 0}^{\infty} \frac{\lambda^\alpha}{\alpha!} e^{-\lambda} K_{\alpha-1} = g(\lambda) \quad \text{(say)}
\]

\[
g'(\lambda) = \frac{d}{d\lambda} (g(\lambda)) = \sum_{\alpha = 0}^{\infty} \frac{1}{\alpha!} [e^{-\lambda} \lambda^{\alpha-1} - e^{-\lambda} \lambda^\alpha]
\]

\[
= \sum_{\alpha = 0}^{\infty} \frac{\lambda^{\alpha-1}}{(\alpha-1)!} - \sum_{\alpha = 0}^{\infty} \frac{\lambda^\alpha}{\alpha!}
\]

\[
= e^{-\lambda} \sum_{\alpha' = 0}^{\infty} \frac{\lambda^{\alpha'}}{\alpha'!} - \sum_{\alpha = 0}^{\infty} \frac{\lambda^\alpha}{\alpha!}
\]

\[
= -e^{-\lambda} \sum_{\alpha = \alpha'}^{\infty} \frac{\lambda^{\alpha}}{\alpha!}
\]

\[
\Rightarrow \int_0^\infty g(\lambda) \, d\lambda = \int_0^\infty -e^{-u} u^k \frac{1}{k!} \, du
\]

\[
\Rightarrow \left[ g(u) \right]_{0}^{\infty} = -\frac{1}{\Gamma(k+1)} \int_0^\infty (e^{-u} u^k) \, du
\]

\[
\Rightarrow g(\lambda) = \frac{1}{\Gamma(k+1)} \int_0^\infty e^{-u} u^k \, du
\]

\[
\therefore \text{distribution function of } P_X \text{ is a decreasing function of } \lambda.
\]

**Note:**

\[
P[X > k] = 1 - P[X \leq k] = \int_0^\infty e^{-u} u^k \, du
\]

\[
= \frac{1}{\Gamma(k+1)} \int_0^\infty e^{-u} u^k \, du
\]

\[
= \frac{1}{\Gamma(k+1)} \left[ \int_0^\infty e^{-u} u^k \, du - \int_0^\infty e^{-u} u^k \, du \right]
\]

\[
= \frac{1}{\Gamma(k+1)} \int_0^\infty e^{-u} u^k \, du
\]

\[
\therefore P[X \leq k] = \sum_{\alpha = 0}^{k} \frac{\lambda^\alpha}{\alpha!} e^{-\lambda} K_{\alpha-1}
\]

\[
= \frac{1}{\Gamma(k+1)} \sum_{\alpha = 0}^{k} \binom{k}{\alpha} \Gamma(k-\alpha+1) \cdot e^{-\lambda} \lambda^\alpha
\]

\[
= \frac{1}{\Gamma(k+1)} \sum_{\alpha = 0}^{k} \binom{k}{\alpha} e^{-\lambda} \lambda^\alpha \int_0^\infty e^{-u} u^{k-\alpha} \, du
\]
\[ F(x) = \frac{1}{\Gamma(k+1)} \int_0^x e^{-(\lambda+u)} \sum_{k=0}^K \frac{(\lambda+u)^k}{k!} \, du \]

\[ = \frac{1}{\Gamma(k+1)} \int_0^x (\lambda+u)^k e^{-(\lambda+u)} \, du \]

\[ = \frac{1}{\Gamma(k+1)} \left[ \int_0^x t^k e^{-t} \, dt \right] \quad \text{[Substitute } t = \lambda + u] \]

\[ = \frac{1}{\Gamma(k+1)} \left[ \int_0^x t^k e^{-t} \, dt \right] \]

\[ = 1 - \frac{\int_0^x t^k e^{-t} \, dt}{\Gamma(k+1)} \]

**Problem 10.** The r.v.'s \( x_1 \) & \( x_2 \) has Poisson distn. with parameters \( \lambda_1 \) & \( \lambda_2 \) respectively \( \lambda_1 > \lambda_2 \), s.t.

\[ P[X_1 \leq k] > P[X_2 \leq k] \]

**Soln.** For any integer \( k \),

Distribution function of \( X \),

\[ F(x) = \frac{1}{\Gamma(k+1)} \int_0^x e^{-u} u^k \, du \]

\[ = 1 - \frac{\int_0^x e^{-u} u^k \, du}{\Gamma(k+1)} \]

So, if \( \lambda \) increases, the corresponding distribution function for any integer \( k \) decreases.

Hence, \( \lambda_1 > \lambda_2 \Rightarrow P[X_1 \leq k] > P[X_2 \leq k] \). \( \text{(Proved)} \)

**Problem 11.** If \( X \sim \text{P}(\lambda) \), then show that

\[ P[X > n] = \frac{\lambda^n}{n!} \]

**Soln.**

\[ P[X > n] = P[X > n-1] = \frac{\int_0^\infty e^{-u} u^{n-1} \, du}{\Gamma(n)} \]

\[ = \frac{n^{n-1} \, du}{\Gamma(n)} \quad \text{[Using } e^{-u} < 1] \]

\[ = \frac{n}{n!} \quad \text{(Proved)} \]
Reproductive property of Poisson Distribution:

A distn. is said to have the reproductive property if

\[ X_1, X_2, \ldots, X_n \] are independently distributed Poisson variables

and \( X_i \sim P(\lambda_i) \) for \( i = 1(1)n \), then if \( S_n = \sum_{i=1}^{n} X_i \sim P \left( \sum_{i=1}^{n} \lambda_i \right) \),

then it is said to have the reproductive property.

**Proof:**

\[ P_{X_i}(t) = e^{-\lambda_i(1-t)} \]

\[ \text{Now, } P_{S_n}(t) = P_{X_1}(t) \cdot P_{X_2}(t) \ldots \cdot P_{X_n}(t) \leftarrow \text{Xi's sum} \]

\[ = e^{-\sum_{i=1}^{n} \lambda_i(1-t)} \]

\[ = e^{-\sum_{i=1}^{n} \lambda_i} \]

\[ \therefore \sum_{i=1}^{n} X_i \sim P \left( \sum_{i=1}^{n} \lambda_i \right) \]

Clearly, the reproductive property is additive.

Binomial distribution conditioning on Poisson variables:

Suppose \( X_1 \) and \( X_2 \) are independently distributed Poisson variables \( P(\lambda_1) \) and \( P(\lambda_2) \), respectively. Then,

\[ \left\{ \frac{X}{X_1 + X_2 = k} \right\} \sim Bin \left( k, \frac{\lambda_1}{\lambda_1 + \lambda_2} \right) \]

i.e., in case the Poisson variables become i.i.d., we get

\[ Bin \left( k, \frac{1}{2} \right) \]

Distribution function of Poisson distn. in terms of the distfn. of a continuous theoretical distribution:

If \( X \sim P(\lambda) \), then,

\[ P[X \leq k] = 1 - \frac{\int_{0}^{\lambda} e^{-\lambda} \lambda^k \, d\lambda}{\Gamma(k+1)} \]

Let,

\[ \frac{\int_{0}^{\lambda} e^{-\lambda} \lambda^k \, d\lambda}{\Gamma(k+1)} = F(\lambda) \]

\( F \) is the distribution function of the standard Gamma distribution with parameter \( (k+1) \). \( F(\lambda) \) is tabulated as \( \frac{\Gamma \left( \frac{\lambda}{k+1} \right)}{\Gamma(k+1)} \) in incomplete Gamma Table.
Problem 12: If \( X \sim P(\lambda) \) and \( Y/X = x \sim Bin(x, \alpha) \), then s.t. \( Y \sim P(\lambda \alpha) \).

Solv.

Hence,

\[
\begin{align*}
\phi(x) &= \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} x \alpha, \\
\phi(y) &= \sum_{y=0}^{\infty} \frac{e^{-\lambda} \lambda^y}{y!} y \alpha - y, \quad y = 0, 1, 2, \ldots, \\
\end{align*}
\]

Hence,

\[
\phi(y) = \sum_{x=y}^{\infty} \phi(y/x) \phi(x).
\]

Now, we can say, \( Y \sim P(\lambda \alpha) \) for \( y = 0, 1, 2, \ldots \).

Hence the proof is complete.
Problem 18. Suppose the no. of eggs laid by an insect have a Poisson distribution with parameter $\lambda$ and the probability of eggs developing in $p$. Show that the no. of eggs surviving has also a Poisson distribution with parameter $\lambda p$, assuming mutual independence of eggs.

Solution. Let us define two random variables:

- $X$ denotes no. of eggs laid by an insect.
- $Y$ denotes no. of insects born.

$$P[Y = y] = \sum_{x=0}^{\infty} P[X = x] P[Y = y | X = x]$$

$$= \sum_{x=0}^{\infty} \frac{e^{-\lambda} \lambda^x}{x!} \left( \frac{\lambda p}{(\lambda p)^{y-1}} \right) e^{-\lambda p} \frac{\lambda^y}{y!} (\lambda p)^{y-y}$$

$$= \frac{e^{-\lambda} \sum_{x=0}^{\infty} \left( \frac{\lambda p}{\lambda p} \right)^{y-y} \frac{(\lambda p)^x}{x!} \frac{\lambda^y}{y!} (\lambda p)^{y-y}}{y!}$$

$$= \frac{e^{-\lambda} \lambda^y}{y!} \frac{\lambda^y}{y!}$$

$$= e^{-\lambda p} (\lambda p)^y$$

$$\Rightarrow Y \sim \text{Poi} (\lambda p)$$
**GEOMETRIC DISTRIBUTION**

**Definition:** A random variable $X$ is said to follow a geometric distribution with parameter $p$ if it has the p.m.f. of the following forms:

- \[ f(x) = pq^x, \quad x = 0, 1, 2, \ldots \quad (\text{Model I}) \]
- \[ f(x) = pq^{x-1}, \quad x = 1, 2, 3, \ldots \quad (\text{Model II}) \]

where \( 0 < p < 1 \) and \( q = 1 - p \).

We denote \( X \sim \text{Geo}(p) \) or \( X \sim \text{Geometric}(p) \).

**Model I**

Figure showing column diagram; features of the distribution are quite evident from the column diagram where we see that the first mass point zero is the most probable value, i.e., the mode of the distribution, and the distribution is positively skewed. \( f(1) \) is maximum.

**Derivation of the p.m.f.:**

**Probability model:** Let us consider a sequence of Bernoulli trials with success probability $p$. Let the trials be repeated till the occurrence of the 1st success. Define:

- \( X \): \# failures preceding the 1st success.
- \( Y \): \# trials required to get the 1st success.

\[ P[X = x] = P[\alpha \text{ failures have occurred before the first success}] = P[F^x S] = pq^x, \quad x = 0, 1, 2, \ldots \quad (\text{Model I}) \quad 0 < p < 1, p + q = 1 \]

\[ P[Y = x] = P[\text{In the first } (x+1) \text{ trials failure has occurred and in the last trial a success have occurred}] = P[F^x S] = pq^x, \quad x = 1, 2, \ldots \quad (\text{Model II}) \quad 0 < p < 1, q = 1 - p \]

i.e. \( X \sim \text{NB}(1, p) \) or \( X \sim \text{Geo}(p) \).

**Note:** An R.V. $X$ that has a geometric distribution is also referred to as a Discrete waiting (time) r.v., since it represents how long (in terms of failures/ trials) one has to wait for the 1st success.

**Example:** Trials of keys required for opening a door.
Mean & Variance:

We know, \( \sum_{x=0}^{\infty} q^x = \frac{1}{1-q} \)

Now, differentiating both sides w.r.t. \( q \),

\[
\frac{d}{dq} \left[ \sum_{x=0}^{\infty} q^x \right] = \frac{d}{dq} \left( \frac{1}{1-q} \right)
\]

\[
\Rightarrow \sum_{x=1}^{\infty} xq^{x-1} = \frac{1}{(1-q)^2} \quad \cdots \cdots \text{(1)}
\]

\[
E(X) = \sum_{x=1}^{\infty} x \cdot pq^x
\]

\[
= \sum_{x=1}^{\infty} x pq^x \quad [\text{Using (1)}]
\]

\[
= \frac{pq}{(1-q)^2} = \frac{q}{p}
\]

Again,

\[
\sum_{x=2}^{\infty} x(x-1)q^{x-2} = \frac{2}{(1-q)^3} \quad \cdots \cdots \text{(2)}
\]

\[
E[X(X-1)] = \sum_{x=0}^{\infty} x(x-1)pq^x = \frac{2pq}{(1-q)^3} = \frac{2q}{p}
\]

\[
V(X) = \frac{2aq}{p^2} + \frac{a}{p} - \frac{q}{p^2} \cdot \frac{a}{p} = \frac{aq}{p^2}
\]

Now, as \( 0 < p < 1 \) and \( p+q = 1 \), so,

\[
\text{Mean} < \text{Variance}
\]

PMF:

\[
P(t) = \sum_{x=0}^{\infty} t^x \cdot pq^x
\]

\[
= \frac{p \sum_{x=0}^{\infty} (at)^x}{(1-at)}
\]
Factorial moment generating function:

\[ P(1+t) = P \left[ 1 - q \left(1 + \frac{t}{1} \right) \right]^{-1} \]
\[ = P \left[ q - \frac{q}{t} \right]^{-1} \]
\[ = \left[ 1 - \frac{q}{p} \cdot \frac{t}{1} \right]^{-1} \]
\[ = \sum_{\alpha=0}^{\infty} \left( \frac{q}{p} \right)^{\alpha} \alpha \]
\[ \Rightarrow \mu[X] = \text{Mean of } X \alpha \]
\[ = \text{Coefficient of } t^{\alpha} \left( \frac{q}{p} \right)^{\alpha} \]
\[ = \alpha! \left( \frac{q}{p} \right)^{\alpha} \]
\[ \Rightarrow \mu_1 = \mu[X] = \frac{q}{p} \]
\[ \mu_2 = \mu_2[X] + \mu_1 = 2 \left( \frac{q}{p} \right) + \frac{q}{p} \]

\[ \text{MGF:} \]
\[ M(t) = \sum_{\alpha=0}^{\infty} e^{\alpha t} \cdot p q^{\alpha} = \sum_{\alpha=0}^{\infty} \left( q e t \right)^{\alpha} = \frac{p}{1 - q e t} \]
\[ = p \left[ 1 - q \left(1 + \frac{t}{1} + \frac{t^2}{2!} + \ldots \right) \right]^{-1} \]
\[ = p \left[ q - \frac{q t}{1} + \frac{q t^2}{2!} + \ldots \right]^{-1} \]
\[ = \left[ 1 - \frac{q}{p} \left( t + \frac{t^2}{2!} + \ldots \right) \right]^{-1} \]
\[ = 1 + \frac{q}{p} \left( t + \frac{t^2}{2!} + \ldots \right) \]
\[ + \left( \frac{q}{p} \right)^{2} \left( t + \frac{t^2}{2!} + \ldots \right) + \ldots \]

\[ \text{Mean} = \text{Coefficient of } t = \frac{q}{p} \]
\[ \mu_2 = \text{Coefficient of } t^{\frac{2}{1}} = \frac{q}{p} + 2 \left( \frac{q}{p} \right)^{2} \]
\[ \Rightarrow \text{Variance} = 2 \left( \frac{q}{p} \right) + \frac{q}{p} - \frac{q}{p} \]
\[ = \frac{q}{p} \]

Factorial moments:

\[ \mu_{[X]} = \text{E}[(X)^{n}] = \sum_{\alpha=0}^{\infty} (X)^{n} \cdot p q^{\alpha} \]
\[ = p q^{n} \left( 1 - q \right)^{-n+1} \cdot n! \]
\[ = n! \left( \frac{q}{p} \right)^{n} \]
**LOSS OF MEMORY PROPERTY**

Geometric distribution lacks memory, i.e., if $X \sim \text{Geo}(p)$, then,

$$P[X > i+j / X > i] = P[X > j], \quad i, j = 0, 1, 2, \ldots$$

Hence, the converse is also true, i.e., the property is a characterisation of Geometric distribution.

And it is called the **Loss of Memory Property** of the geometric distribution. \[\text{weseu'11}\]

**Proof:**

If \(i \leq j \rightarrow \) $X \sim \text{Geo}(p)$ then,

$$P[X > i+j / X > i] = \frac{P[X > j]}{P[X > i]} \quad ; \quad i, j = 0, 1, 2, \ldots$$

**Ans:**

$$P[X > i] = \sum_{x=i}^{\infty} pq^x = p \frac{q^i}{1-q}$$

$$\therefore P[X > i+j] = q^{i+j}$$

$$\therefore P[X > i+j / X > i] = \frac{P[X > i+j]}{P[X > i]} = \frac{q^{i+j}}{q^i} = q^j$$

$$\therefore P[X > i+j] = P[X > i] P[X > j] \quad \text{(Proved)}$$

Only if \(i \leq j \rightarrow \) $P[X > i+j / X > i] = P[X > j]$\]

Then, \(\rightarrow X \sim \text{Geo}(p)\)

**Ans:**

$$r_i = P[X > i]$$

$$\therefore r_{i+j} = P[X > i+j] = P[X > i] P[X > j] = p_i \cdot p_j$$

$$r_{i+1} = p_i, \quad r_i = r_{i-1} = \cdots = p_i, \quad r_i = (i-1)$$

$$p_i = P[X > 1] = 1 - P[X = 0]$$

$$\therefore P[X = 0] = 1 - p_i$$

$$p_2 = P[X > 2] = 1 - P[X = 0] - P[X = 1]$$

$$\therefore P[X = 1] = 1 - p_2 - (1 - p_i)$$

$$= 1 - p_i^2 - 1 + p_1 = p_1 (1 - p_i)$$
\[ P_x = P[x \geq 3] = 1 - P[x = 0] - P[x = 1] - P[x = 2] \]
\[ = 1 - (1-P_1)(1-P_1)P_1^3 \]
\[ = P_1^3(1-P_1) \]

Let it be true for \( i = m-1 \),

i.e., \( P_{m+1} = P_1^{m+1} \)

and \( P[ x = m-1 ] = P_1^{m-1} (1-P_1) \)

\[ \therefore P[ x > m+1 ] = 1 - \sum_{j=m+1}^{\infty} P[ x = j ] \]
\[ = P_1^{m+1}(1-P_1) \]

\( \therefore X \sim \text{Geo}(P) \).

\[ \text{Example} \Rightarrow \text{As a crude example, if the arrival of a bus at a specific bus stop in each minute (unit of time) is considered as a Bernoulli trial with a constant probability } P \text{ the probability that a person has to wait at least for 5 minutes, remains independent of whether he has already waited for 3 minutes or is just coming.} \]

\[ \text{Problem 1.} \quad \text{let } X \sim \text{Geo}(P), \text{then show that } \therefore f(x) = pq^x, x=0,1,\ldots \]
\[ P[ x > i+j ] = P[ x > i ] P[ x > j ] \]

\[ \text{Proof:} \]
\[ P[ x > i+j ] = \sum_{x=i+j}^{\infty} pq^x = q^{i+j+1} \]

Similarly, \( P[ x > i+j ] = q^{i+j+1} \)

\[ \therefore P[ x > i+j ] = P[ x > i+j ] P[ x > j ] \]

\[ \because P[ x > i+j ] = P[ x > i+j ] P[ x > j ] = P[ x > i+j ] / P[ x > j ] \]

\[ \therefore \text{Note:} \quad \text{in Problem 1, } \]
\[ \therefore \text{it says that the Geometric distribution has no memory, that is, the information of no successes in } i \text{ trials is forgotten in subsequent calculations.} \]
\[ \therefore \text{the converse of the theorem is also true & discussed in the next page.} \]
**Problem 2.** Let \( X \) be a non-negative integer valued random variable satisfying the equation

\[
P[X > i + j] = p[X > i] p[X > j] \quad \forall \, i, j = 1, 2, \ldots, \infty.
\]

Then show that \( X \) must have a geometric distribution.

**Proof:** Let \( q_i = P[X > i] \), \( i = 0, 1, \ldots \).

Then \( q_{i+j} = P[X > i] \cdot P[X > j] = P[X > i] \cdot P[X > j] = P[X > i] \cdot P[X > j] = \ldots = q_{i} \cdot q_{j-1} \)

Taking \( j = 1, \) we get \( q_{i+1} = q_{0} \cdot q_{i} = q_{0} \cdot q_{i-1} = q_{0}^{2} \cdot q_{i-2} = \ldots = q_{0}^{i+1} \cdot q_{0} = q_{0}^{i+2} \)

\[
q_{i} = q_{0}^{i+1}
\]

\[
\therefore \, P[X = i] = P[X > i - 1] - P[X > i]
\]

\[
= q_{i-1} - q_{i}
\]

\[
= q_{0}^{i} - q_{0}^{i+1}
\]

\[
= q_{0}^{i} (1 - q_{0})
\]

\[
\therefore \, X \sim \text{Geo}(q_{0}).
\]

**Problem 3.** Let \( X \) & \( Y \) be i.i.d. R.V.'s, and let

\[
P[X = k] = p_{k} > 0, \quad k = 0, 1, 2, \ldots
\]

if \( P[X = t | X + Y = t] = \frac{P[X = t - 1 | X + Y = t]}{1 + t} \), \( t > 0 \)

then \( X \) & \( Y \) are geometric R.V.'s.

**Solution:**

\[
P[X = t | X + Y = t] = P[X = t - 1 | X + Y = t]
\]

\[
P[X = t, X + Y = t] = \frac{P[X = t - 1, X + Y = t]}{P[X + Y = t]}
\]

\[
P[X = t, Y = 0] = P[X = t - 1, Y = 1]
\]

\[
P[X = t | P[Y = 0] = P[X = t - 1, P[Y = 1]]
\]

\[
\frac{P_{t - 1}}{t} = \frac{P_{0}}{P_{0}}
\]

Hence,

\[
P_{t} = \frac{P_{t - 1}}{P_{t - 1}} x \frac{P_{t - 2}}{P_{t - 2}} x \ldots x \frac{P_{0}}{P_{0}} x P_{0}
\]

\[
= \frac{P_{1}}{P_{0}} x \frac{P_{1}}{P_{0}} x \ldots x \frac{P_{1}}{P_{0}} x P_{0}
\]

\[
= \left( \frac{P_{1}}{P_{0}} \right)^{t} \cdot P_{0}
\]

Since, \( \sum_{t=0}^{\infty} P_{t} = 1 \), we must have \( \frac{P_{1}}{P_{0}} < 1 \).

Moreover,

\[
P_{0} \sum_{t=0}^{\infty} \left( \frac{P_{1}}{P_{0}} \right)^{t} = 1 \Rightarrow P_{0} \cdot \frac{1}{1 - \left( \frac{P_{1}}{P_{0}} \right)} = 1
\]

\[
\Rightarrow \frac{P_{0}}{P_{0}} = \left( 1 - \frac{P_{1}}{P_{0}} \right)
\]

\[
\therefore \, P_{t} = (1 - \frac{P_{1}}{P_{0}}) \cdot P_{0}
\]

\[
\therefore \, \text{The proof is complete.}
\]
Another way: we have,

\[ P[X = t \mid X+Y = t] = \frac{P_t P_0}{\sum_{k=0}^{t+1} P_k P_{t-k}} = \frac{1}{t+1}. \]

and

\[ P[X = t-1 \mid X+Y = t] = \frac{1}{t+1}. \]

It follows that

\[ \frac{P_t}{P_{t-1}} = \frac{P_t}{P_0}. \]

---

**Problem 1.** Let \( X \) be a non-negative, integer-valued RV satisfying

\[ P[X > m+1 \mid X > m] = P[X > 1] \]

for any non-negative integer \( m \), then \( X \) must have a geometric distribution.

**Proof:** Let the PMF of \( X \) be written as

\[ P[X = k] = P_k \quad k=0, 1, 2, \ldots. \]

Then

\[ P[X > m] = \sum_{k=m}^{\infty} P_k \]

and

\[ P[X > m] = \sum_{k=m+1}^{\infty} P_k = q_{m+1}, \text{ say}. \]

Thus,

\[ q_{m+1} = q_m q_0, \]

where \( q_0 = P[X > 0] = P_1 + P_2 + \cdots = 1 - P_0 \). It follows that

\[ q_k = (1 - P_0)^{k+1}, \]

and hence

\[ P_k = q_{k-1} - q_k = (1 - P_0)^k P_0, \text{ as asserted}. \]

---

**Interpretation of Loss of memory property:** Probability that more than \( m+n \) trials will be required before the first success given that there have been already more than \( m \) failures is equal to the unconditional probability of at least \( n \) trials are needed before the 1st success. Therefore, the information of no of successes in \((m+n)\) trials is forgotten in subsequent calculation.
Problem 5. Express negative binomial distribution as a sum of a number of geometric distribution.

**Solution:**

$x$ can be written as, $X = X_1 + X_2 + \cdots + X_n$, where $X_i$ = No. of failures preceding the $i$th success after having $(i-1)$ success.

$X_i$'s are i.i.d. geometric random variables having p.m.f.

$$P(X_i = x) = pq^x \text{ for } x = 0, 1, \ldots$$

P.G.F of $X$ is given by,

$$P_X(t) = \left[ P_{X_i}(t) \right] = \frac{1}{1 - qt} = \sum_{x=0}^{\infty} t^x pq^x = (1 - qt)^{-p}$$

$$= \sum_{x=0}^{\infty} \frac{(-qt)^x}{x!} \cdot p^x (-qt)^x$$

$$P(X = x) = \text{coefficient of } t^x$$

$$= \binom{x + n - 1}{n - 1} p^n q^x$$

$$E(X) = E\left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} \frac{a}{i} = \frac{np}{p} = \frac{r}{p}$$

$$V(X) = V\left( \sum_{i=1}^{n} X_i \right) = \sum_{i=1}^{n} V(X_i) + 0 \quad [\text{covariance term vanishes due to independence}]$$

$$= \sum_{i=1}^{n} \frac{a^2}{i^2} = \frac{r^2}{p^2}$$
HYPERGEOMETRIC & WAITING TIME HYPERGEOMETRIC
(INVERSE HYPERGEOMETRIC) DISTRIBUTION:

Probability Model: Consider a box containing \( N \) objects of which \( M \) are of one kind (say, type A) and \( N-M \) are of other kind (say, type B).

Suppose \( n \) objects are drawn at random from the box.

\[ X = \text{No. of type A objects drawn in the sample,} \]

Then the p.m.f. of \( X \) is

\[
P[X = x] = \begin{cases} 
\binom{M}{x} \binom{N-M}{n-x} \left( \frac{N}{n} \right) & \text{if } x \in \left[ \max(0, n+M-N), \min(M, n) \right] \\
0 & \text{otherwise}
\end{cases}
\]

Note, \( 0 \leq x \leq M \)
\( 0 \leq n-x \leq N-M \)
\( \Rightarrow m+M-N \leq x \leq n \)
\( \Rightarrow x \in \max(0, n+M-N), \ldots, \min(M, n) \).

If the R.V. \( X \) follows a hypergeometric distribution with parameters \( n, N, M \),
we write \( X \sim H(n, N, M) \).

Let,
\[
p = \frac{M}{N} = \text{Proportion of type A objects in the box.}
\]
\[
q = \frac{N-M}{N} = 1-p
\]
\[
f(x) = \left( \frac{Np}{x} \right) \left( \frac{Nq}{n-x} \right) \left( \frac{N}{n} \right) \]
\[
\text{if } x = 0, 1, \ldots, \min(m, n) \text{, } p, q > 0, \]
\[
p + q = 1
\]

We write, \( X \sim H(n, N, p) \)

In practice \( n \) is so chosen that \( n \leq \min(Np, Nq) \).
Then the mass points of \( X \) are \( 0, 1, 2, \ldots, n \).
Factorial moments:  
\[ f(x) = \frac{(\frac{Np}{n})^x}{\binom{N}{n}} \]

\[ M[k] = E(X^k) = \sum_{x=k}^{n} \binom{n}{x} \frac{(1+x)^{Np}}{(x+k)!(Np-x)!} \left( \binom{N}{n-x} \right) \]

\[ = \frac{1}{\binom{N}{n}} \sum_{x=k}^{n} \frac{(Np)!}{(Np-x)!} \binom{Np-x}{x-k} \binom{n-x}{N-1-n+k} \]

\[ = \frac{1}{\binom{N}{n}} \binom{Np}{k} \sum_{x=k}^{n} \binom{n-x}{N-1-n+k} \binom{Np-k}{x-k} \binom{n-x}{N-1-n+k} \]

\[ = \frac{1}{\binom{N}{n}} \binom{Np}{k} \sum_{x=0}^{n-k} \binom{n-x}{N-1-n+k} \binom{Np-k}{x-k} \binom{n-x}{N-1-n+k} \]

Note: \((1+t)^{Np-x} (1+t)^{n-x} = (1+t)^{N-k}\)

\[ \Rightarrow \left\{ \ldots \cdot \left( \frac{Np}{x} \right)^{t+x} \ldots \right\} \left\{ \ldots \cdot \left( \frac{N}{n-x} \right)^{t+n-k-x} \ldots \right\} \]

Collecting coefficient of \(t^{n-k}\) from both sides we get,

\[ \sum_{x=0}^{n-k} \binom{Np-k}{x} \binom{N}{n-x} = \binom{N-k}{n-k} \]

\[ \therefore M[k] = \frac{\binom{Np}{k}}{\binom{N}{n}} \binom{N-k}{n-k} \]

\[ \therefore M[1] = E(X) = \frac{\binom{Np}{n}}{\binom{N}{n}} = np \]

\[ \therefore M[2] = E[X(X-1)] = \frac{NP(NP-1)}{\binom{N}{n}} \binom{N-2}{n-2} = \frac{(NP-1)(n-1)np}{(N-1)} \]

\[ \therefore M_2 = var(X) = \frac{P(NP-1)n(n-1)}{N-1} + np(n-1)^2 = \frac{npa - N-n}{N-1} \]

\[ \therefore M_3 = npa(2-p) \frac{(N-n)(N-2n)}{(N-1)(N-2)} \]

\[ \therefore M_4 = \frac{npa(N-n)}{(N-1)(N-2)(N-3)} \left[ N(N+1) - Gn(N-n) + Pq - \frac{N(N-2)(N-1)}{6} n(N-n) \right] \]
Mode of the Hypergeometric Distribution:

\[ f(x) = \binom{Np}{x} \binom{N-np}{n-x} / \binom{N}{n} = \frac{n!}{(n-x)!(x-1)!} \binom{N}{n} \]

\[ f(x-1) = \frac{n!}{(n-x)(x-1)!} \binom{N}{n} \]

\[ f(x) \geq f(x-1) \] according as \[ \frac{(n-x+1)(n-x+2)}{x(n+1)} \geq 1 \]

On, \( Np = Np + Np = x + Np \approx Np \approx 0 \)
On, \( Np + Np + n \approx (N+2)x \)
On, \( (Np+1)(n+1) \approx (N+2)x \)
\[ \Rightarrow x \approx \frac{(Np+1)(n+1)}{(N+2)} \]

\[ \Rightarrow \alpha \text{ has mode at } x = \left[ \frac{(Np+1)(n+1)}{(N+2)} \right] \text{ if } \alpha \text{ is not an integer.} \]

If \( \alpha \) be an integer, then two modes are at \( \frac{(Np+1)(n+1)}{(N+2)} - 1 \) and \( \frac{(Np+1)(n+1)}{(N+2)} \).

Binomial Approximation to Hypergeometric Distribution:

If \( X \sim H(N, n, p) \) and \( N \gg n \), then \( \frac{n}{N} \) can be ignored, i.e. \( \frac{i}{n} \rightarrow 0 \) as \( N \rightarrow \infty \) and \( n = \text{fixed} \). Then the Hypergeometric distribution will converge to a binomial distribution with parameters \( n \) & \( p \).

**Proof:** For some fixed \( \alpha \)

\[ P[X=x] = \binom{Np}{x} \binom{N-np}{n-x} / \binom{N}{n} \]

\[ = \frac{n!}{(n-x)!(x-1)!} \binom{N}{n} \]

\[ = \frac{(n-1)!}{(n-x)!(x-1)!} \binom{N}{n} \]

\[ = \frac{N!}{(n-1)!(x-1)!} \binom{N}{n} \]

\[ = \frac{n!}{(n-x)!(x-1)!} \binom{N}{n} \]

\[ = \binom{n}{x} p^x (1-p)^{n-x} \]

As \( N \rightarrow \infty \), \( P \frac{1}{N} \rightarrow p \) \( \forall j = 0 \) \( N \rightarrow \infty \)

**Interpretation:** If \( N \rightarrow \infty \), the population becomes an infinite population. From an infinite population, drawing samples WR and WOR are practically same, since one may have a sampling unit repeated under WOR. Then \( P[X=x] \) under WOR tends to \( P[X=x] \) under WR as \( N \rightarrow \infty \).
Mean Deviation about Mean

\[ n \text{ items } = a + b \]

\[ p = \frac{a}{N}, \quad p + q = \frac{a + b}{N} = 1 \]

\[ q = \frac{b}{N}, \quad 0 \leq p \leq 1. \]

\( n \) elements are drawn without replacement

\[ P[X = \alpha] = \left( \binom{N}{\alpha} \binom{n - \alpha}{n - \nu} \right) / \binom{N}{n} \]

Mean deviation about mean is defined as

\[ E[|X - np|] \]

\[ = \frac{2}{\binom{N}{n}} \sum_{\alpha = n_0}^{n} (x - np) \binom{N}{x} \binom{n - \alpha}{n - x}, \text{ where } n_0 = \lceil np + 0.5 \rceil \]

\[ = \frac{2}{\binom{N}{n}} \sum_{\alpha = n_0}^{n} \left[ \alpha (p + q) - np \right] \binom{N}{x} \binom{n - \alpha}{n - x} \]

\[ = \frac{2}{\binom{N}{n}} \sum_{\alpha = n_0}^{n} \left[ \alpha q - (n - x) p \right] \binom{N}{x} \binom{n - \alpha}{n - x} \]

\[ = \frac{2}{\binom{N}{n}} \left[ a \sum_{\alpha = n_0}^{n} N_0 \binom{N - 1}{x - 1} \left( \frac{N - 1}{n - x - 1} \right)^a \left( \frac{N - 1}{n - x - 1} \right)^b - p \sum_{\alpha = n_0}^{n} (n - x) \binom{n - \alpha}{n - x} \binom{N - 1}{x} \left( \frac{N - 1}{n - x - 1} \right)^a \left( \frac{N - 1}{n - x - 1} \right)^b \right] \]

\[ = \frac{2 N p q}{\binom{N}{n}} \left[ \sum_{\alpha = n_0}^{n - 1} \binom{N - 1}{x - 1} \binom{N - 1}{n - x - 1} + \sum_{\alpha = n_0}^{n-1} \binom{N - 1}{x - 1} \binom{N - 1}{n - x - 1} \right] \]

\[ \quad - \sum_{\alpha = n_0}^{n-1} \binom{N - 1}{x - 1} \binom{N - 1}{n - x - 1} \left( \frac{N - 1}{n - x - 1} \right)^a \left( \frac{N - 1}{n - x - 1} \right)^b \]

\[ = \frac{2 N p q}{\binom{N}{n}} \left[ \sum_{\alpha = n_0}^{n - 1} \binom{N - 1}{x - 1} \binom{N - 1}{n - x - 1} \right] \]

\[ = \frac{2 N p q}{\binom{N}{n}} \left( \frac{N - 1}{n_0 - 1} \right) \left( \frac{N - 1}{n - n_0} \right) \]

\[ = \frac{2 N p q}{\binom{N}{n}} \frac{n_0}{N} \frac{\binom{N}{n_0}}{\binom{N}{n_0 - 1}} \frac{\binom{N - n + n_0}{n_0}}{\binom{N}{n_0}} \]

\[ = \frac{2 n_0 (N_0 - m + n_0)}{\binom{N}{n_0}} P[X = n_0] \]
Problem 1. If \( X \sim H(n, n, p) \), then find the distribution of \( (n - X) \).

\[ P[X = k] = \binom{n}{k} \left( \frac{np}{n} \right)^k \left( \frac{n(1-p)}{n} \right)^{n-k} \]

\[ P[X = n-k] = \binom{n}{n-k} \left( \frac{np}{n} \right)^{n-k} \left( \frac{n(1-p)}{n} \right)^k \]

\[ P[\gamma = \alpha] = \binom{n}{\alpha} \left( \frac{np}{n} \right)^\alpha \left( \frac{n(1-p)}{n} \right)^{n-\alpha} \]

Let, \( \gamma = n - X \).

\[ P[\gamma = \alpha] = \binom{n}{\alpha} \left( \frac{np}{n} \right)^\alpha \left( \frac{n(1-p)}{n} \right)^{n-\alpha}, \text{ where } \gamma \sim H(n, n, p). \]

Problem 2. If \( X \sim H(n, n, p) \), then show that \( \gamma(X) \leq \frac{n}{4} \).

\[ \gamma(X) = \frac{np(1-n)}{N-1} \]

\[ \frac{d\gamma(X)}{dp} = \frac{n(n-n)}{N-1} \left[ q - p \right] = 0 \]

\[ p = \frac{1}{2} \]

\[ \frac{d^2\gamma(X)}{dp^2} = \frac{n(n-n)}{N-1} \left[-1-1 \right] < 0 \]

\( \gamma(X) \) is maximum at \( p = \frac{1}{2} \).

\[ \gamma(X) \leq \frac{n}{4} \cdot \frac{n-n}{N-1}, \text{ since } \frac{n-n}{N-1} \leq 1 \text{ as } n \geq 1 \]

\[ \gamma(X) \leq \frac{n}{4} \cdot \frac{n-n}{N-1} \text{ for } n \geq 1 \]
**Discrete Uniform Distribution**:

Let \( X \) be a r.v. that takes only the values \( x_1, x_2, \ldots, x_N \). Then \( X \) is said to have a uniform distribution over the set \( \{ x_1, x_2, \ldots, x_N \} \) if \( P[X = x_i] = \text{constant} \times (1/N) \).

Now, 
\[
1 = P[X = \omega] = P \left[ \bigcup_{i=1}^{N} \{ X = x_i \} \right] \\
= \sum_{i=1}^{N} P[X = x_i] \\
= \sum_{i=1}^{N} \frac{1}{N} \\
= N \frac{1}{N} = 1, 
\]

\( \Rightarrow k = \frac{1}{N} \).

**Definition:**

A r.v. \( X \) is said to have a uniform distribution over \( \{ x_1, x_2, \ldots, x_N \} \) if its PMF is:

\[
f_X(x) = \begin{cases} \frac{1}{N} & \text{if } x = x_i \quad \forall i = 1(1)N \\ 0 & \text{otherwise} \end{cases}
\]

**Moments:**

\[
E(X) = \sum_{i=1}^{N} x_i P[X = x_i] \\
= \frac{1}{N} \sum_{i=1}^{N} x_i \\
= \bar{x} \quad \text{(say)}
\]

\[
\text{Var}(X) = E(X^2) - E^2(X) \\
= \frac{1}{N} \sum_{i=1}^{N} x_i^2 - \bar{x}^2 \\
= \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 = s^2 \quad \text{(say)}.
\]

In descriptive statistics, if we assume that the sampling values are uniformly distributed, then

Sample Mean = \( \bar{x} \) = \( \frac{1}{N} \sum_{i=1}^{N} x_i \), and

\[
S^2 = \frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2 = \text{Variance}.
\]
**Rectangular Distribution or Uniform Distribution:**

An absolutely continuous random variable $X$ defined over $[a, b]$, $-\infty < a < b < \infty$ is said to follow uniform distribution with parameter $a, b$ if its pdf is given by,

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

We will write $X \sim U[a, b]$ if $X$ has a uniform distribution on $[a, b]$.

This distribution is also called a rectangular distribution since the area under the graph of $f$ between $a$ and $b$ is rectangular. It is also called Rectangular Distribution.

$X \sim U[a, b]$

or

$X \sim R[a, b]$

The end points $a$ or $b$ or both may be excluded. Clearly,

$$\int_{a}^{b} f(x) dx = 1.$$ 

**Distribution Function:**

The df of $X$ is given by,

$$F(x) = \begin{cases} 
0, & \text{if } x < a, \\
\frac{x-a}{b-a}, & \text{if } a < x < b, \\
1, & \text{if } x > b, 
\end{cases}$$

**Expectation and Variance:**

$$E(X^k) = \int_{a}^{b} x^k f(x) dx = \frac{1}{b-a} \int_{a}^{b} x^k dx, \quad k > 0 \text{ is an integer,}$$

$$= \frac{1}{b-a} \left[ \frac{x^{k+1}}{k+1} \right]_{a}^{b},$$

Putting $k=1$, $E(X) = \frac{b+a}{2}$.

Putting $k=2$, $E(X^2) = \frac{(b-a)^2 + ab + a^2}{3}$.

$$\therefore \text{Var}(X) = E(X^2) - E(X)^2 = \frac{(b-a)^2}{12}.$$
Moment Generating Function:

\[ M_X(t) = E(e^{tx}) = \int_{b-a}^{e^{tx} dx} \frac{e^{tb} - e^{ta}}{t(b-a)} \]

\[ = \frac{1}{t(b-a)} \left[ \sum_{j=0}^{\infty} \frac{(tb)^j}{j!} - \sum_{j=0}^{\infty} \frac{(ta)^j}{j!} \right] \]

\[ = \frac{1}{t(b-a)} \sum_{j=0}^{\infty} \frac{j!(tb)^j - (ta)^j}{j!} \]

\[ = \frac{1}{t(b-a)} \sum_{j=0}^{\infty} \frac{t^j(b^j - a^j)}{j!} \]

\[ \therefore \mu' = E(X) = \frac{b-a}{2(b-a)} = \frac{b+a}{2} \]

\[ \therefore \mu_2' = E(X^2) = \frac{b^3-a^3}{3(b-a)} = \frac{b+ab+a^2}{3} \]

\[ Y(X) = \frac{b+ab+a^2}{3} - \left( \frac{a+b}{2} \right)^2 = \frac{(b-a)^2}{12} \]

S.D. = \[ \frac{b-a}{\sqrt{12}} \]

NOTE: The distribution is trivially symmetric about \[ \frac{a+b}{2} \]

Theorem 1: Probability Integral Transformation

Let \( X \) be a continuous R.V. having D.F. \( F \). Then \( F(X) \) has the uniform distribution on \([0,1]\).

Proof:- Let, \( Y = F(X) \)

Then the D.F of \( Y \) is given by

\[ F_Y(y) = P[Y \leq y] = P[F(X) \leq y] = P[X \leq F^{-1}(y)] = F[F^{-1}(y)] \]

[as \( F(x) \) is a monotone non-decreasing function]

\[ = \left\{ \begin{array}{ll}
0 & \text{if } y \leq 0 \\
1 & \text{if } 0 < y < 1 \\
1 & \text{if } y \geq 1 
\end{array} \right. \]

P.D.F of \( Y = \frac{d}{dy} [F_Y(y)] = 1 \)

\[ Y = F(X) \sim U(0,1) \]

Note that, \( E[Y] = \frac{1}{2} \), \( V[Y] = \frac{1}{12} \).
NOTE: The fact can be used to draw random observations from the theoretical distribution of $X$.

Here at first we choose $B$ digits random numbers and put a decimal point before the first digit. Let us denote such a quantity by $p$, clearly $p$ is a realization from $U(0,1)$ dist., now to obtain $x$ we equate $F(x) = p$ and solve for $x$.

**Theorem:** If $F$ be any DF, and let $X$ be a $U[0,1]$ RV. Then there exists a function $h$ such that $h(X)$ has DF $F$, i.e.,

$$P\{h(X) \leq \alpha\} = F(\alpha) \quad \text{for all } \alpha, \beta \in (-\infty, \infty).$$

**Proof:** If $F$ is the DF of a discrete RV $Y$, let

$$P\{Y = y_k\} = p_k, \quad k = 1, 2, \ldots$$

Define $h$ as follows:

$$h(x) = \begin{cases} y_1 & \text{if } 0 \leq x < p_1, \\ y_2 & \text{if } p_1 \leq x < p_1 + p_2, \\ \vdots & \vdots \\ y_k & \text{if } p_{k-1} \leq x < \sum_{i=1}^{k} p_i, \quad k = 1, 2, \ldots. \end{cases}$$

Then

$$P\{h(X) = y_k\} = P\{0 \leq x < p_1\} = p_1$$

and in general,

$$P\{h(X) = y_k\} = p_k, \quad k = 1, 2, \ldots$$

Thus $h(X)$ is a discrete RV with DF $F$.

If $F$ is continuous and strictly increasing, $F^{-1}$ is well defined, and we take

$$P\{h(X) \leq \alpha\} = P\{F^{-1}(X) \leq \alpha\} = P\{X \leq F(\alpha)\}$$

as asserted.

In general, define

$$F^{-1}(y) = \inf \{ x : F(x) \geq y \}$$

and let $h(X) = F^{-1}(X)$, then we have

$$F^{-1}(X) \leq \alpha \iff \{ x : F(x) \leq \alpha \} \subseteq \{ y \geq F(x) \}.$$
Problem: Why is the distribution called rectangular distribution?

Ans:
\[ f(x) = \begin{cases} \frac{1}{b-a}, & a < x < b \\ 0, & \text{otherwise} \end{cases} \]

The graph of the pdf \( f(x) \) looks like a rectangular and it uniform distribution over \([a, b]\) is also known as Rectangular Distribution.

\[ X \sim U[a, b] \]

Remark:

Let \( X \sim U[-\frac{1}{2}, \frac{1}{2}] \)

Then the b.d.f. of \( X \) is \( f(x) = \begin{cases} 2, & \frac{1}{2} < x < 1 \\ 0, & \text{otherwise} \end{cases} \)

Note that the b.d.f. \( f(x) \) takes values greater than unity.

In case of Uniform distribution,
\[ \frac{1}{b-a} = \text{length of the interval} \]

CDF:
\[ F_X(x) = \int_{-\infty}^{x} f_X(t) \, dt = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x < b \\ 1, & x \geq b \end{cases} \]
**Gamma Distribution:**

- **Definition:** An absolutely continuous random variable \( X \) defined over \((0, \infty)\) is said to follow Gamma distribution with parameters \( \alpha \) and \( \beta \) if its pdf is given by:

\[
 f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} \quad \text{for } x > 0 ; \alpha, \beta > 0
\]

where, \( \Gamma(\beta) \) is of the form, \( \Gamma(\beta) = \int_0^\infty e^{-t} t^{\beta-1} dt \) is the Gamma function.

Thus, we write \( X \sim \text{Gi} (\alpha, \beta) \).

- **Derivation of the pdf:** A probability law can easily be obtained from the following improper integral:

\[
 f(x) = \int_{-\infty}^{\infty} e^{-x^2} dx
\]

Here, the improper integral converges iff \( \alpha > 0, \beta > 0 \).

Now,

\[
 \int_0^{\infty} e^{-\alpha x} x^{\beta-1} dx = \frac{1}{\alpha^\beta} \int_0^{\infty} e^{-2} \left( \frac{x}{\alpha} \right)^{\beta-1} dx \quad \text{[let, } \alpha^\beta = 2]\]

\[
 = \frac{1}{\alpha^\beta} \int_0^{\infty} e^{-2} \frac{x^{\beta-1}}{\alpha^\beta} dx
\]

\[
 = \frac{\Gamma(\beta)}{\alpha^\beta}
\]

Now,

\[
 \int_0^{\infty} x^{\beta-1} \left( e^{-\alpha x}, x^{\beta-1} \right) dx = 1 \Rightarrow \text{It is verified that it's a pdf.}
\]

We can define:

\[
 f(x) = \frac{\alpha^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\alpha x} \quad \text{for } x > 0, \alpha > 0, \beta > 0
\]

Since the pdf is obtained from Gamma integral, the distribution is referred to as Gamma Distribution.

We denote the distribution by \( \text{Gi} (\alpha, \beta) \) or \( \Gamma(\alpha, \beta) \).

Clearly \( \gamma(\alpha, 1) \) is an exponential distribution with mean \( \alpha \).

**Note:** \( \gamma(1, \beta) \) is known as a standard Gamma Distribution, and we denote the distribution by \( \gamma(\beta) \).
Properties:

1. Gamma distribution is bell-shaped, positively skewed depending on the choice of the leptokurtic distribution. It may be slightly skewed and leptokurtic or, it may be significantly skewed and leptokurtic. Hence, the term of the Gamma density can be thought of slight on significant departure from natural density.

\[ \alpha = 0.5 \]
\[ \alpha = 1 \]
\[ \alpha = 2 \]
\[ P = 0.8 \]
\[ P = 2 \]
\[ P = 4 \]
\[ P = 8 \]

\[ f(x) \]
\[ P = 1 \]
\[ P = 2 \]
\[ P = 8 \]

Gamma density function:

Gamma densities (\( \alpha = 1 \))
2. Additive Property of Gamma distribution:

Suppose $X_1, \ldots, X_k$ are independently distributed gamma variables where

$$X_i \sim \Gamma(\alpha_i, \lambda_i) \quad i = 1, \ldots, k.$$ 

Define, $S_k = \sum_{i=1}^{k} X_i$ then $S_k \sim \Gamma(\alpha, \lambda), \sum_{i=1}^{k} \lambda_i.$

**Reproducible property:**

$$M_{S_k}(t) = \prod_{i=1}^{k} M_{X_i}(t) = \frac{\Gamma\left(\alpha + \sum_{i=1}^{k} \lambda_i\right)}{\Gamma(\alpha)} \left(1 - \frac{t}{\alpha}\right)^{-\sum_{i=1}^{k} \lambda_i}.$$

$$\therefore S_k \sim \Gamma\left(\alpha, \sum_{i=1}^{k} \lambda_i\right).$$

Note that the distribution function of gamma distribution can be expressed explicitly in terms of that of a suitable Poisson distribution.

**Moments:**

$$\mu_\nu' = E(X^\nu)$$

$$= \int_0^\infty x^{\nu p_1 - 1} e^{-x\alpha} \alpha^\nu p^\nu \, dx$$

$$= \frac{\Gamma(\nu + p)}{\Gamma(\nu) \alpha^\nu}.$$ 

Thus in case of standard Gamma distribution where $\alpha = 1,$

then Mean = Variance, \(\ast\)

$$\begin{align*}
\alpha & = 1, \\
\mu_1' & = E(X) = \frac{(p+1)}{\alpha} = \frac{p}{\alpha}.
\end{align*}$$

$$\begin{align*}
\alpha & = 2, \\
\mu_2' & = E(X^2) = \frac{(p+1)(p+3)}{\alpha^2} = \frac{p(p+1)}{\alpha^2}.
\end{align*}$$

$$\nu = 1, \quad \nu' = \frac{p}{\alpha}.$$ 

Mode of the distribution:

$$\mu_1 = \frac{\alpha P \cdot e^{-x\alpha}}{\Gamma(\nu)}.$$ 

$$\frac{d}{dx} [\mu_1] = \frac{\alpha^p}{\Gamma(\nu)} \left[ \alpha^{p-1} \cdot (-\alpha) \cdot e^{-x\alpha} + \alpha^\nu \cdot (p-1)\alpha^{p-2} \right]$$

$$= \frac{\alpha^p}{\Gamma(\nu)} \cdot \alpha^{p-2} \cdot e^{-x\alpha} \left[ -\alpha \alpha + p - 1 \right] = 0$$

$$\therefore \alpha = \frac{p-1}{\alpha} \text{ when } p > 1; \quad \mu_1 > 0 \text{ when } \alpha < \frac{p-1}{\alpha}; \quad \mu_1 < 0 \text{ when } \alpha > \frac{p-1}{\alpha}.$$
Moment Generating Function:

\[ M_G(t) = M_X(t) = E(e^{tX}) \]

\[ = \int e^{tx} \frac{\alpha^p}{\Gamma(p)} e^{-\alpha x} x^{p-1} dx \]

\[ = \frac{\alpha^p}{\Gamma(p)} \int e^{-(\alpha-t)x} x^{p-1} dx \]

\[ = \frac{\alpha^p}{\Gamma(p)} \left( 1 - \frac{t}{\alpha} \right)^{-p} \quad \text{when} \ |t| < \alpha, \]

\[ = \sum_{k=0}^{\infty} \frac{(-p)^k}{k!} \left( \frac{-t}{\alpha} \right)^k \]

\[ \mu'_n = \text{Coefficient of } \frac{t^n}{n!} = (-p)^n \frac{(-1)^n}{\alpha^n} \]

\[ \mu'_1 = \frac{p}{\alpha} = \text{Mean.} \]

\[ \mu'_2 = \frac{p(p+1)}{\alpha^2} \]

\[ \mu'_2 = \frac{\mu}{\alpha^2} \]

\[ \mu_2 = \frac{p(p+1)}{\alpha^2} - \frac{p}{\alpha^2} = \frac{p}{\alpha^2} \Rightarrow \text{Var}(X) \]

Cumulant Generating Function:

\[ K(t) = \ln M(t) \]

\[ = -p \ln \left( 1 - \frac{t}{\alpha} \right) \quad |t| < \alpha \]

\[ = p \sum_{n=1}^{\infty} \frac{\left( \frac{t}{\alpha} \right)^n}{n!} = p \sum_{n=1}^{\infty} \frac{t^n}{n!} \left( \frac{1}{\alpha} \right)^n \]

\[ K_{10} = \frac{p(10-1)!}{\alpha^{10}} = \text{Coefficient of } \frac{t^{10}}{10!} \]

Mean = \[ \frac{p}{\alpha} \]; variance = \[ \frac{p}{\alpha^2} = \mu_2 \]

\[ \mu_3 = \mu_8 = \frac{2p}{\alpha^3}; \quad \mu_4 = 3k_2 + k_4 = \frac{6p}{\alpha^4} + \frac{3p^2}{\alpha^4} = \frac{3p(2+p)}{\alpha^4} \]

\[ \beta_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2p}{\sqrt{3} \alpha^3 \sqrt{p}} = \frac{2}{\sqrt{p}} > 0 \]

Hence, the distribution is positively skewed.

\[ \beta_2 = \frac{\mu_4}{\mu_2^2} = 3 = \frac{6p}{p} > 0 \]

The distribution is leptokurtic.
Distribution function of a Gamma Random Variable:

Let \( X \sim \text{Gamma}(\alpha, \beta) \).

The distribution function of \( X \) is given by,

\[
F(x) = P[X \leq k] = \int_0^k \frac{\alpha^\beta}{\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx.
\]

Now,

\[
I_p = \int_0^k \frac{\alpha^\beta}{\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx = \frac{\alpha^\beta}{\Gamma(\alpha)} \left[ -\frac{\beta}{\alpha} e^{-\frac{x}{\beta}} x^{\alpha-1} \right]^k_0
\]

\[
+ (p-1) \int_0^k \frac{\alpha^\beta}{\Gamma(\alpha)} e^{-\frac{x}{\beta}} x^{\alpha-2} \, dx
\]

\[
= -\frac{\alpha^\beta}{\Gamma(\alpha)} e^{-\frac{x}{\beta}} x^{\alpha-1} \bigg|_0^k + \frac{\alpha^\beta}{\Gamma(\alpha)} \int_0^k e^{-\frac{x}{\beta}} x^{\alpha-2} \, dx
\]

\[
= -\frac{\alpha^\beta}{\Gamma(\alpha)} e^{-\frac{x}{\beta}} x^{\alpha-1} + I_{p-1}
\]

\[
= \frac{(\alpha^\beta)^{p-1} e^{-\frac{x}{\beta}} - (\alpha^\beta)^{p-2} e^{-\frac{x}{\beta}} - \cdots - (\alpha^\beta) e^{-\frac{x}{\beta}} + 1}{\Gamma(\alpha)}
\]

Thus,

\[
P[X > k] = \sum_{i=0}^{p-1} \frac{(\alpha^\beta)^i e^{-\frac{x}{\beta}}}{\Gamma(\alpha)} = P[Y \leq p-1], \text{ where } Y \sim \text{Beta}(\alpha, \beta).
\]

Problem 1: If a random variable has a pdf \( f(x) = \frac{1}{\alpha \beta} x^{\alpha-1} e^{-\frac{x}{\beta}} \),

\( \alpha > 0 \) and \( \alpha, \beta > 0 \), show that

\[
E\{g(x)(x-\alpha\beta)\} = \beta \cdot E\{xg(x)\}, \text{ provided both the expectations exist}.
\]

Ans: \[
E\{g(x)(x-\alpha\beta)\} = \int g(x)(x-\alpha\beta) \cdot \frac{1}{\alpha \beta} x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx
\]

\[
= \frac{1}{\alpha \beta} \left[ g(x) \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} \right]_0^{+\infty} - \alpha \beta E[g(x)]
\]

\[
= \frac{1}{\alpha \beta} \int g(x) \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx - \alpha \beta E[g(x)]
\]

\[
= \frac{1}{\alpha \beta} \left[ g(x) \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} \right]_0^{+\infty} + \beta \int e^{-\frac{x}{\beta}} [\alpha x^{\alpha-1} g(x) + x^{\alpha-1} g'(x)] \, dx
\]

\[
= \frac{1}{\alpha \beta} \left[ g(x) \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} \right]_0^{+\infty} + \beta \int e^{-\frac{x}{\beta}} [\alpha x^{\alpha-1} g(x) + x^{\alpha-1} g'(x)] \, dx
\]

\[
= \frac{1}{\alpha \beta} \left[ g(x) \cdot x^{\alpha-1} e^{-\frac{x}{\beta}} \right]_0^{+\infty} + \beta \int e^{-\frac{x}{\beta}} [\alpha x^{\alpha-1} g(x) + x^{\alpha-1} g'(x)] \, dx
\]
\[
\frac{1}{E(x)} = \frac{1}{E(x^{-1})} = \frac{1}{\mu_1^{-1}} = \frac{1}{(\alpha-1)\Gamma(p)} = \frac{\alpha}{p-1}
\]

**Problem 2.** Evaluate the Harmonic Mean (HM) of the Geometric distribution.

**Solution:**

\[
H_M = \frac{1}{E(\frac{1}{X})} = \frac{1}{E(X^{-1})} = \frac{1}{\mu_1^{-1}} = \frac{\Gamma(p)}{\Gamma(p-1)} = \frac{\alpha}{p-1}
\]

**Another Form of the PDF of Gamma Distribution:**

\[
f(x) = \begin{cases} \frac{1}{\theta^n \Gamma(n)} e^{-x/\theta} x^{n-1} & \text{if } 0 < x < \infty \\ 0 & \text{otherwise} \end{cases}
\]

**Derivation of the P.D.F.:** Considering the improper integral, \( \int e^{-x/\theta} x^{n-1} dx \)

It converges if \( \theta > 0, n > 0 \), \( X \sim \text{Gamma}(\theta, n) \)

Now, \( \int x^{n-1} e^{-x/\theta} dx = \int e^{z/\theta} z^{n-1} \cdot \theta^n dz \)

\[
z = \frac{x}{\theta} \Rightarrow dx = \theta dz
\]

\[
\int x^{n-1} e^{-x/\theta} dx = \frac{\theta^n}{\theta^n} \int z^{n-1} e^{-z/\theta} dz = 1 \Rightarrow \text{It is verified that it's}
\]

The above integral converges for \( \theta > 0, n > 0 \)

So, \( X \sim \Gamma(\theta, n) \) on Gamma \( (\theta, n) \)

**Moments:**

\[
\mu_n = E(X^n) = \int_0^\infty \frac{x^n}{\theta^n \Gamma(n)} e^{-x/\theta} x^{n-1} dx
\]

\[
= \int_0^\infty e^{-x/\theta} x^{n-1} dx = \frac{\theta^n}{\theta^n} \frac{\Gamma(n+1)}{\Gamma(n+1)} = \frac{\theta^n}{\theta^n} \frac{\Gamma(n+1)}{\Gamma(n+1)} = \frac{\theta^n}{\theta^n}
\]

**Mean:**

\[
\frac{\theta_{n+1}}{\theta_n} = \theta_n, \text{Variance } = \mu_2 - \mu_1^2 = \frac{\theta_{n+2}}{\theta_n} - (\theta_n)^2 = \theta^2 n
\]
\section*{Beta Distribution:}

- Beta Distribution of 1st & 2nd kind: \rightarrow The idea of probability density can easily be developed from following improper integral:

\[ \int_0^1 x^{a-1} (1-x)^{b-1} \, dx = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)} \quad a, b > 0 \]

\[ \int_0^{\infty} \frac{x^{a-1}}{(1+x)^{a+b}} \, dx = \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)} \quad a, b > 0 \]

- (i) & (ii), respectively, lead to the following probability density:

\[ f_1(x) = \frac{1}{\beta(a,b)} x^{a-1} (1-x)^{b-1} I_x(0,1) \quad a, b > 0 \quad \cdots \quad (i) \]

\[ f_2(x) = \frac{x^{a-1}}{(1+x)^{a+b}} I_x(0,\infty) \quad a, b > 0 \quad \cdots \quad (ii) \]

Clearly, \( a > 0, b > 0 \)

\[ f_1(x) \geq 0 \quad \forall x \quad \& \quad f_2(x) \geq 0 \quad \forall x \quad \& \quad \int_0^\infty f_1(x) \, dx = 1 \quad \& \quad \int_0^\infty f_2(x) \, dx = 1 \]

The probability distribution having pdf \( f_1(x) \) is called Beta distribution of 1st kind, we denote this dist by \( \beta_1(a,b) \), and the probability distribution with pdf \( f_2(x) \) is called Beta distribution of 2nd kind, we denote this dist by \( \beta_2(a,b) \).

- Depending on the parameters \( a, b \), Beta distribution of the 1st kind may be uniform, symmetric bell-shaped, symmetric distribution, skewed bell-shaped, positively skewed J-shaped, U-shaped, negatively skewed bell-shaped, and negatively skewed J-shaped, \( \cdots \). Depending upon the parameters, some of the moment of Beta distribution of 2nd kind do not exist. Hence, MGF of Beta distribution does not exist. For the existence of \( \ell \)th order normal moments, we can easily obtain as follows:

\[ \mu_n = \int_0^1 x^{a+n-1} (1+x)^{-a-b} \, dx = \frac{\Gamma(a+n) \Gamma(1+b)}{\Gamma(1+a+n) \Gamma(b)} \]

\[ \mu_n = \frac{1}{\beta(a,b)} \beta(a+n, b+n) \]

[Provided, \( a > -n \) and \( b > n \)]

\[ \text{If we consider the integer \( \ell \) moments, then the } \ell \text{th order moment exists if } \ell \text{ exceeds } n. \]

\[ \text{Note: if } x \sim \beta_2(a,b) \implies \frac{x}{1+x} \sim \beta_1(a,b) \]

\[ \text{or, conversely, if } x \sim \beta_1(a,b) \implies \frac{x}{1-x} \sim \beta_2(a,b) \]

\[ \int_0^1 x \frac{\beta_2(a,b)}{\beta(a,b)} \, dx = 1 \implies \frac{1}{\beta(a,b)} \int_0^1 x \beta_2(a,b) \, dx = 1 \implies \frac{2}{1+b} = 1 \]

\[ \int_0^1 x \frac{\beta_2(a,b)}{\beta(a,b)} \, dx = 1 \implies \frac{1}{\beta(a,b)} \int_0^1 x \beta_2(a,b) \, dx = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]

\[ \frac{2}{1+b} = 1 \implies b = 1 \]
Beta Distribution of 1st kind:

Definition: An absolutely continuous random variable $X$ defined over $[0, 1]$ is said to follow a Beta distribution with parameters $a$ and $b$ if its p.d.f. is given by:

$$f(x) = \begin{cases} \frac{1}{\beta(a,b)} x^{a-1} (1-x)^{b-1}, & 0 < x < 1, \ a > 0, \ b > 0 \\ 0 & \text{otherwise} \end{cases}$$

Notation: $X \sim \beta(a,b)$ or simply $\beta(a,b)$.

Moments:

$$\mu_n = E(X^n) = \int_0^1 x^n \frac{1}{\beta(a,b)} x^{a-1} (1-x)^{b-1} \, dx$$

$$= \frac{\beta(a+n, b)}{\beta(a,b)} = \frac{\Gamma(a+b+n)}{\Gamma(a+b) \Gamma(a+n) \Gamma(b+n)} = \frac{(a+b-1)! (a+b-1)!}{(a+b+n-1)! (a-1)!} \frac{a^n b^n}{\Gamma(a) \Gamma(b)}$$

Putting $n = 1$, $E(X) = \frac{a}{(a+b)} \frac{(a+b-1)!}{(a+b)! (a-1)!} = \frac{a}{a+b}$.

Putting $n = 2$, $E(X^2) = \frac{(a+1)}{(a+b+1)} \frac{(a+b-1)!}{(a+b+1)! (a-1)!} = \frac{a(a+1)}{(a+b)(a+b+1)}$.

Variance ($\text{Var}(X)$):

$$\text{Var}(X) = \frac{a(a+1)}{(a+b)(a+b+1)} - \frac{a^2}{(a+b)^2}$$

$$= \frac{ab}{(a+b)^2(a+b+1)} \left[ a(a+1)(a+b) - a^2(a+b+1) \right]$$

**Problem 1.** If $X \sim \beta(a,b)$ then show that $V(X) \leq \frac{1}{4}$.

**Solution:**

$$V(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

$$\frac{1}{4} - V(X) = \frac{(a+b)^2(a+b+1) - 4ab}{4(a+b)^2(a+b+1)}$$

$$= \frac{(a+b)^2(a+b+1) - (a+b)(a+b)}{4(a+b)^2(a+b+1)}$$

$$= \frac{(a+b)^2(a+b+1)}{4(a+b)^2(a+b+1)} = \frac{1}{4}$$

Thus, $V(X) \leq \frac{1}{4}$.
Problem 2. If \( x \sim B(a, b) \); s.t. AM > HM,

\[
E(x) = \frac{a}{a+b} \quad E\left(\frac{1}{x}\right) = \frac{a-2}{(a+b-2)!(a-1)!} = \frac{a+b-1}{a-1}.
\]

\[
\therefore \frac{1}{E\left(\frac{1}{x}\right)} = \frac{a-1}{a+b-1}, \quad a > 1
\]

\[
\therefore E(x) - \frac{1}{E\left(\frac{1}{x}\right)} = \frac{a}{a+b} - \frac{a-1}{a+b-1} = \frac{b}{(a+b)(a+b-1)} > 0 \quad \text{[} a > 1 \text{]}
\]

\[
\therefore E(x) > \frac{1}{E\left(\frac{1}{x}\right)} \quad \text{AM > HM}
\]

Problem 3. If \( x \sim B(a, b) \), find the geometric mean.

Solution:

Let the geometric mean be \( g \).

\[
\ln g = \int \ln x \cdot f(x) \, dx = \int \ln x \cdot \frac{x^{a-1}(1-x)^{b-1}}{\beta(a, b)} \, dx
\]

Now, \( \beta(a, b) = \int_0^1 x^{a-1}(1-x)^{b-1} \, dx \)

\[
\frac{\partial \beta(a, b)}{\partial a} = \int_0^1 x^{a-1}(1-x)^{b-1} \ln x \, dx
\]

\[
\therefore \ln g = \frac{1}{\beta(a, b)} \cdot \frac{\partial \beta(a, b)}{\partial a}
\]

\[
= \frac{\partial}{\partial a} \ln \left( \frac{\Gamma(a) \Gamma(b)}{\Gamma(a+b)} \right)
\]

\[
= \frac{\partial}{\partial a} \left[ \ln \Gamma(a) + \ln \Gamma(b) - \ln \Gamma(a+b) \right]
\]

\[
= \frac{\partial}{\partial a} \ln \Gamma(a) - \frac{\partial}{\partial a} \ln \Gamma(a+b) \quad (\text{Ans})
\]
Problem 4. If \( X \sim \beta(a, b) \) then show that:

\[
E \left[ \left( \frac{b - (a-1)(1-x)}{x} \right) g(x) \right] = E \left[ (1-x) g'(x) \right]
\]

Ans.:
\[
E \left[ \left( \frac{b - (a-1)(1-x)}{x} \right) g(x) \right] = \int_0^1 \left( \frac{b - (a-1)(1-x)}{x} \right) g(x) f(x) \, dx
\]
\[
= b \cdot E[g(x)] - \int_0^1 (a-1) g(x) \cdot x^{a-2} (1-x)^{b-1} \cdot \frac{1}{B(a,b)} \, dx
\]
\[
= b \cdot E[g(x)] - \left[ \frac{a-1}{(a-1)} \cdot \frac{x^{a-1}}{B(a,b)} \right]_0^1 \cdot \int_0^1 x^{a-2} (1-x)^{b-1} \, dx
\]
\[
= b \cdot E[g(x)] + \int_0^1 g'(x) (1-x)^{b-1} \cdot \frac{1}{B(a,b)} \cdot x^{a-1} (1-x)^{b-1} \, dx
\]
\[
= E \left[ (1-x) g'(x) \right]
\]

MGF of Beta Distribution:

The MGF of \( X \) is given by:

\[
M_X(t) = E(e^{tx}) = E \left( \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} \right)
\]

Since \( 0 < x < 1 \), the MGF converges.

Now, \( \sum_{n=0}^{\infty} \frac{(tx)^n}{n!} < \sum_{n=0}^{\infty} \frac{|t|^n}{n!} = e^{|t|} \)

\( e^{|t|} \) is a convergent series.

And hence, \( M_X(t) = \sum_{n=0}^{\infty} \frac{E(X^n)}{n!} \)

\[
= \sum_{n=0}^{\infty} \frac{t^n}{n!} \cdot \frac{B(a+n, b)}{B(a, b)}
\]

\[
= \sum_{n=0}^{\infty} \frac{t^n}{n!} \left( \prod_{k=0}^{n-1} \frac{a+k}{a+b+k} \right)
\]

Which can't be expressed in a compact form.
EXPERIMENTAL DISTRIBUTION:

One parameter Exponential Distribution:

Derivation of the pdf: Suppose we are given the distribution of life in hours of thousand dry cells of a particular brand. Hence the observed distribution will be,

\[
\text{Relative frequency density} \uparrow
\]

\[
\text{life } \rightarrow
\]

Note that, here the frequency density,

\[
f(x) = \frac{e^{-\Theta x}}{\Theta}, \quad \text{where } \Theta > 0
\]

determines the law explicitly. Therefore, it is reasonable to assume a probability density function,

\[
f(x) = e^{-\Theta x} I_{x}(0,\infty)
\]

If we regard average life to be \( \mu \) then,

\[
k \int e^{-\Theta x} \, dx = \mu
\]

\[
\Rightarrow \kappa \int e^{-\Theta x} \, dx = \mu
\]

\[
\Rightarrow \frac{1}{\Theta} = \frac{\kappa}{\mu}
\]

\[
\Rightarrow \Theta = \frac{\mu}{\kappa}
\]

\[
\text{It's the p.d.f. of one parameter Exponential distribution. We write, } X \sim \text{E}(\mu)
\]

\[
f(x) = \frac{1}{\mu}e^{-x/\mu}, \quad \text{where } 0 < x < \infty, \mu > 0
\]

Note: Here we see that the distribution is J-shaped, positively skewed, thus we don't find any measure of kurtosis. We also see that mean (= \( \mu \)) of the distribution characterises the location as well as the spread of the distribution since in case of Exponential distribution,

Definition: An absolutely continuous random variable is said to follow exponential distribution if the p.d.f. is of the form,

\[
f(x) = \frac{1}{\mu}e^{-x/\mu}, \quad 0 < x < \infty, \mu > 0
\]

We write: \( X \sim \text{E}(0,\mu) \) or \( X \sim \text{E}(\mu) \).
Moment generating function: \( M(t) = MGF = \frac{1}{\mu} \int_0^\infty e^{tx} - \frac{x}{\mu} \, dx = E(e^{tx}) \)

\[ = \frac{1}{\mu} \int_0^\infty e^{-\frac{x}{\mu}(1-\mu t)} \, dx \]

[exists if \( |t| < \frac{1}{\mu} \)]

\[ = \frac{1}{\mu} \frac{\mu}{1-\mu t} \]

\[ = (1-\mu t)^{-1} \]

\[ = \sum_{j=0}^\infty \mu^j t^j \quad \text{[exists if \( |t| < \frac{1}{\mu} \)]} \]

\[ \therefore \mu_j = j! \frac{\mu^j}{j!} \]

\[ \therefore \mu_j = \text{coefficient of } \frac{t^j}{j!} \text{ in the expansion of } M(t) \]

\[ \therefore \mu_1 = \text{mean } = \mu \]

\[ \therefore \mu_2 = 2 \mu^2 \quad \therefore \gamma(x) = 2\mu^2 - \mu^2 = \mu^2 \]

Cumulant Generating function:

\[ k(t) = \ln M(t) \]

\[ = \ln (1-\mu t)^{-1} \]

\[ = -\ln (1-\mu t) \]

\[ = \sum_{j=1}^\infty \frac{\mu^j t^j}{j} \]

\[ \therefore k_j = \text{coefficient of } \frac{t^j}{j!} \]

\[ = \mu^j (j-1)! \]

\[ k_1 = \mu = \text{mean} \]

\[ k_2 = \mu^2 = \text{variance, s.d. } = \mu \]

\[ k_3 = \mu^3 \]

\[ k_4 = 6/\mu^4 \]

\[ \mu_4 = k_4 + 3k_2^2 = 9\mu^4 \]

CDF:

\[ F_X(x) = \int_{-\infty}^x f_X(t) \, dt = \int_{-\infty}^x \frac{1}{\mu} e^{-\frac{t}{\mu}} \, dt = 1 - e^{-x/\mu}, \quad x > 0 \]

\[ \therefore F_X(x) = x \quad \text{for } x < 0 \]
Exponential Distribution lacks memory:

**Loss of memory property:**

- If $X$ is exponential with mean $\mu$ then $P(X > x + y \mid X > x) = P(X > y)$

**Proof:**

- If $X \sim E(\mu)$ then $P(X > x + y \mid X > x) = P(X > y)$

**Ans:**

\[
P(X > y) = \int_{x+y}^{\infty} \frac{e^{-x/y}}{y} \, dx
\]

\[
= \frac{1}{\mu} \left[ e^{-x/\mu} \right]_{x+y}^{\infty}
\]

\[
= \frac{1}{\mu} \left[ e^{-x/\mu} - e^{-(x+y)/\mu} \right]
\]

\[
= e^{-y/\mu}
\]

\[
P(X > x + y) = e^{-(x+y)/\mu};
\]

\[
P(X > x) = e^{-x/\mu};
\]

\[
P(X > x + y) = \frac{e^{-(x+y)/\mu}}{e^{-x/\mu}} = e^{-y/\mu} = P(X > y)
\]

$\Rightarrow P(X > x + y \mid X > x) = P(X > y)$.

*Note:* Exponential distribution is the only continuous distribution which lacks memory.

Only if part:

- If $P(X > x + y) = P(X > x)P(X > y)$, then $X \sim E(\lambda)$.

**Ans:**

\[
P(X > x + y + z) = P(X > x)P(X > y + z) = P(X > x)P(X > y)P(X > z)
\]

Let $P(X > x) = e^{-x}$.

*It could be written that,*

\[
\phi(t) = \left[ \sum_{n=0}^{\infty} \frac{t^n}{n!} \right] = \sum_{n=0}^{\infty} \frac{t^n}{n!} \phi(t) = \phi(t)
\]

\[
\phi(\pm t) = \phi(\mp t)
\]

\[
\phi'(t) = \phi'(t)
\]

\[
\phi''(t) = \phi''(t)
\]

\[
\phi(t) = \frac{d}{dt} \phi(t)
\]

\[
\phi(t) = e^{-t}
\]

\[
\phi(t) = \int f(x) \, dx \quad \text{where} \quad f(x) = e^{-x}
\]

**Differentiating both sides we get,**

\[
0 = -f(x) = -xe^{-x} \Rightarrow f(x) = xe^{-x}
\]

\[
x \sim E\left(\frac{1}{2}\right) \quad \text{where mean} = \frac{1}{2}.\]
Note: \( f(x) = \frac{1}{\mu} e^{-x/\mu} \)

\[ \Phi(1) = \int_0^\infty \frac{1}{\mu} e^{-x/\mu} \, dx = \frac{1}{\mu} \left[ -\mu e^{-x/\mu} \right]_0^\infty = \left[ e^{-1/\mu} \right] \]

\[ \ln \Phi(1) = -\frac{1}{\mu} \]

\[ \frac{1}{\mu} = \frac{1}{\ln \Phi(1)} = \frac{1}{\lambda} = \text{mean} \]

2. If \( X \) be a non-negative continuous random variable satisfying the loss of memory property, then \( X \) must be an exponential random variable. This is a characterisation of exponential distribution.

### Mean Deviation about mean:

\[ MD_\mu (X) = \frac{2}{\mu} \int_0^\infty (x-\mu) \cdot e^{-x/\mu} \, dx \]

\[ = \frac{2}{\mu} \int_0^\infty (\alpha-x/\mu) \cdot e^{-x/\mu} \, dx \]

\[ = \frac{2}{\mu} \int_0^\infty t \cdot e^{-(t+\mu)/\mu} \, dt \]

\[ = \frac{2}{\mu} \int_0^\infty t \cdot e^{-t/\mu} \cdot e^{-t/\mu} \, dt \]

\[ = \frac{2}{\mu} \int_0^\infty t \cdot e^{-t/\mu} \, dt \]

\[ = \frac{2\mu}{e} \]

\[ \therefore \quad \text{s.d.} \quad \text{of} \quad X \quad \equiv \quad 2\frac{\mu}{e} \]

**Definition:** A r.v. \( X \) is defined to have an (negative) exponential distribution with parameter \( \mu \) if its pdf is given by

\[ f(x, \mu) = \begin{cases} \frac{1}{\mu} e^{-x/\mu}, & 0 < x < \infty, \mu > 0 \\ 0, & \text{otherwise} \end{cases} \]

The DF is

\[ F_X(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-x/\mu}, & 0 \leq x \end{cases} \]
Distribution of minimum of a set of independently distributed exponential variables:

Suppose \( x_1, x_2, \ldots, x_n \) are independently distributed exponential variables.

Where, \( x_i \sim \text{exponential with mean } \theta_i \)

\[
X(y) = \min \{x_1, x_2, \ldots, x_n\}
\]

\[
P[X(y) \leq x] = 1 - P[X(y) > x]
\]

\[
= 1 - P[x_1, x_2, \ldots, x_n > x]
\]

\[
= 1 - \prod_{i=1}^{n} P[x_i > x]
\]

\[
= 1 - \prod_{i=1}^{n} \left[ \int_{x}^{\infty} \frac{1}{\theta_i} e^{-u/\theta_i} du \right]
\]

\[
= 1 - \prod_{i=1}^{n} e^{-x/\theta_i}
\]

\[
= \prod_{i=1}^{n} \frac{1}{\theta_i} e^{-x/\theta_i}
\]

\[
= \prod_{i=1}^{n} \frac{1}{\theta_i} e^{-x/\theta_i}
\]

Thus \( X(y) \) has exponential distribution with mean \( \frac{1}{\sum_{i=1}^{n} \frac{1}{\theta_i}} \).

If \( x_i \)'s are i.i.d. exponential variables with mean unity then

\[
X(y) \sim \text{exponential with mean } \frac{1}{n}
\]

**Problem:** If \( x \sim \text{E}(0, \theta) \) then show that, \( P[X > r + s | X > r] = P[X > s] \), \( r, s \in \mathbb{R}^+ \).

**Ans:**

\[
P[X \leq r] = 1 - e^{-r/\theta}
\]

\[
P[X > r] = e^{-r/\theta}
\]

\[
P[X > r + s] = e^{-r/\theta}
\]

\[
P[X > r + s] = e^{-r/\theta}
\]

\[
P[X > r + s | X > r] = P[X > s]
\]

\[
P[X > r + s | X > r] = P[X > s]
\]
Problem 2: If \( x \sim \mathcal{N}(0, \sigma) \), derive a recursion relation connecting central moments.

Solution:
Mean = 0,
\[
\mu_r = E[(x - \theta)^r] = \int_0^\infty (x - \theta)^r \frac{x}{\theta} e^{-x/\theta} dx
\]
\[
= \frac{1}{\theta} \int_0^\infty (x - \theta)^r e^{-x/\theta} dx
\]
Differentiating w.r.t. \( \theta \), we get:
\[
\frac{d\mu_r}{d\theta} = -\frac{r}{\theta} \int_0^\infty (x - \theta)^{r-1} e^{-x/\theta} dx - \frac{r}{\theta} \int_0^\infty (x - \theta)^{r-1} e^{-x/\theta} dx
\]
\[
+ \frac{1}{\theta^2} \int_0^\infty (x - \theta)^r e^{-x/\theta} dx
\]
\[
= -r \mu_{r-1} + \frac{1}{\theta^2} \int_0^\infty (x - \theta)^{r-1} e^{-x/\theta} dx
\]
\[
= -r \mu_{r-1} + \frac{1}{\theta^2} \mu_{r+1}
\]
\[
\therefore \mu_{r+1} = \theta^r \left[ \frac{d\mu_r}{d\theta} + r \mu_{r-1} \right]
\]
Putting \( r = 1 \), \( \mu_2 = \theta^1 \cdot [1 + 0] = \theta^2 \)
\( r = 2 \), \( \mu_3 = \theta^2 \cdot [2 + 2 \times 0] = 2 \theta^3 \)
\( r = 3 \), \( \mu_A = \theta^3 \cdot [6 \theta + 3 \times 0] = 9 \theta^4 \)

\[ x \]

[Further calculations]

Scanned by CamScanner  Ctanujit Classes  Page No. 204
• Truncated Exponential / Shifted Exponential Distribution:

**Derivation of the pdf:** Suppose, \( X \) is exponentially distributed with mean \( \theta \) and we are to obtain the distribution of \( X \) given that \( X > \alpha \), the p.d.f. of the truncated distribution is,

\[
g(x) = \frac{1}{\theta} \cdot e^{-x/\theta} \cdot I_x(\alpha, \infty)
\]

\[
= \frac{1}{\theta} \cdot e^{-x/\theta} \cdot I_x(\alpha, \infty)
\]

Choosing \( \theta = b (> 0), \alpha = a, \) then, p.d.f. is

\[
f(x) = \frac{1}{b} \cdot e^{-\frac{x-a}{b}} ; a < x < \infty ; b > 0
\]

**Definition:** An absolutely continuous random variable \( X \) defined over \( [a, \infty) \) where \( a \in \mathbb{R} \) is said to follow exponential distribution with parameter \( \lambda \) and \( b \) if its pdf is given by,

\[
f(x) = \begin{cases} \frac{1}{b} \cdot e^{-\frac{x-a}{b}} ; & a < x < \infty \quad b > 0 \\ 0 & \text{else} \end{cases}
\]

We write \( X \sim E(\lambda, b) \).

**Particular case:** When \( a = 0, b = \mu = \text{mean (}> 0) \), then

\[
f(x) = \frac{1}{\mu} \cdot e^{-x/\mu} ; 0 < x < \infty
\]

which is the p.d.f. of one parameter exponential distribution.

\( X \sim E(0, \mu) \).

**Moments:** The \( n \)th order raw moment of \( X \) about \( a \) is given by

\[
\mu'_n (a) = E((X-a)^n) = \int_a^\infty (x-a)^n \frac{1}{b} \cdot e^{-\frac{x-a}{b}} dx
\]

Let, \( x-a = z, dx = dz \)

\[
= \frac{1}{b} \int_0^\infty z^n \cdot e^{-z/\theta} \ dz
\]

\[
= b^\theta \cdot \Gamma(n+1)
\]

\[
= b^\theta \cdot \frac{(n+1)!}{\theta^n}
\]

* It is important to note that shifted exponential dist. does not have the "lack of memory property."
Cumulant Generating Function:

\[ k(t) = \ln M_X(t) \]
\[ = \ln \frac{e^{bt}}{1 - bt} \]
\[ = at - \ln(1 - bt) \]
\[ = at + \sum_{j=2}^{\infty} \frac{(bt)^j}{j} + bt \]
\[ = (a+b)t + \sum_{j=2}^{\infty} \frac{(bt)^j}{j!} \times (j-1)! \]
\[ \therefore k_1 = E(X) = \mu = a+b = \text{mean} \]
\[ k_2 = \text{variance} = \mu_2 = b^2 \]
\[ k_3 = \mu_3 = 2b^3 \]
\[ k_4 = G_4 \]
\[ \therefore \mu_4 = k_4 + 3k_2 \]
\[ = 6b^4 + 3b^4 \]
\[ = 9b^4 \]

Measures of Skewness:

\[ \gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2b^3}{b^3} = 2 \]
\[ \therefore \text{The distribution is positively skewed.} \]

Measures of Kurtosis:

\[ \gamma_2 = \frac{\mu_4}{\mu_2^2} - 3 = \frac{9b^4}{b^4} - 3 = 6 \]
\[ \therefore \text{The distribution is leptokurtic.} \]

Mode of the distribution:

\[ f(\alpha) = \frac{1}{b} e^{-\left(\frac{\alpha-a}{b}\right)} \]
\[ \therefore \alpha \leq \alpha < \infty, b > 0 \]
\[ \therefore \frac{df}{d\alpha} = -\frac{1}{b} e^{-\left(\frac{\alpha-a}{b}\right)} < 0 \]
\[ \therefore f(\alpha) \text{ is a decreasing function of } \alpha. \]
\[ \therefore f(\alpha) \text{ is maximum when } \alpha \text{ is minimum.} \]
\[ \therefore \text{The mode of the distribution is at } \alpha = a. \]
Mean Deviation about mean:

If \( X \sim E(a, b) \) then \( \mu = a + b \).

\[
MD_\mu(x) = E |X - \mu|
\]

\[
= \int_0^\infty |x-\mu| f(x) dx
\]

\[
= 2 \int_0^\infty (x-\mu) f(x) dx
\]

\[
= 2 \int_0^\infty (x-a-b) \frac{1}{b} e^{-(x-a-b)/b} dx
\]

\[
= 2 \int_0^\infty (b-x) e^{-z} dz
\]

\[
= 2b \int_0^\infty z e^{-z} dz - 2b \int_0^\infty e^{-2z} dz
\]

\[
= 2b \left[ -z e^{-z} \right]_1^\infty + \int_1^\infty e^{-z} dz - 2b \left[ e^{-z} \right]_1^\infty
\]

\[
= 2b \left[ e^{-1} + e^{-1} \right] - 2b \left[ e^{-1} \right]
\]

\[
= 2be^{-1}
\]

Quantiles:

Let \( \xi_p \) denotes the \( p \)th quantile of \( X \).

\[
F(\xi_p) = p \quad \forall \quad p \in (0, 1)
\]

Now, \( F(x) = \frac{1}{b} \int_a^x e^{-\frac{t-a}{b}} dt \)

\[
= \frac{1}{b} \int_0^{\frac{x-a}{b}} e^{-z} b dz
\]

\[
= \left[ e^{-z} \right]_0^{\frac{x-a}{b}}
\]

\[
= 1 - e^{-\frac{x-a}{b}}
\]

\[
\therefore F(\xi_p) = p \quad \Rightarrow \quad 1 - e^{-\frac{\xi_p-a}{b}} = p
\]

\[
\therefore e^{-\frac{(\xi_p-a)}{b}} = 1-p
\]

\[
\therefore -\frac{\xi_p-a}{b} = \ln(1-p)
\]

\[
\therefore \xi_p = a - b \ln(1-p)
\]
1st quartile at \( p = \frac{1}{4} \):
\[
\frac{\ln (1/4)}{\ln 2} = a - \frac{\ln 3}{\ln 2} = a + b \ln \frac{3}{4}
\]

2nd quartile at \( p = \frac{1}{2} \):
\[
\frac{\ln (1/2)}{\ln 2} = a + b \ln 2 = \text{Median}
\]

3rd quartile at \( p = \frac{3}{4} \):
\[
\frac{\ln (3/4)}{\ln 2} = a + b \ln 3
\]

4. **Quantile Deviation:**

\[
\text{Q.D.} = \frac{\frac{3}{4} - \frac{1}{4}}{2} = \frac{\ln 1 - \ln 3/4}{2} = \frac{\ln 3}{2}
\]

**Note:** Hence, \( \text{Q.D.} = \frac{\ln 3}{2b} = \frac{\ln 3}{2} \).

**Problem 1:** If \( X \sim E(a, b) \) then \( \text{S.D.} = \text{mean} > \text{median} > \text{mode} \).

**Ans:** If \( X \sim E(a, b) \) then we have already calculated,

- \( \text{mean} = a + b \), \( \text{median} = a + b \ln 2 \), \( \text{mode} = a \).

- as \( b > 0 \), so \( a + b > a \Rightarrow \text{mean} > \text{mode} \).----(1)

- We know, \( \ln 2 < 1 \)

- \( a + b \ln 2 < a + b \Rightarrow \text{mean} > \text{median} \).----(2)

- \( a + b \ln 2 > a \) as \( b > 0 \), \( \ln 2 > 0 \).

- \( \Rightarrow \text{median} > \text{mode} \).----(3)

- Combining (1), (2), (3), \( \therefore \text{Mean} > \text{median} > \text{mode} \).

**Problem 2:** If \( X \sim E(0, \frac{1}{b}) \), find the mean deviation of \( X \) about its median.

**Ans:** Median = \( a + b \ln 2 \) when \( X \sim E(a, b) \).

Here, \( X \sim E(0, \frac{1}{b}) \), so median = \( \frac{1}{b} \ln 2 = \frac{\ln 2}{b} \).

\[
\text{MD}_{\frac{1}{b}}(X) = E \left| X - \frac{\ln 2}{b} \right| = \frac{e}{-\frac{e}{b}} = \int_{-\infty}^{\infty} (e^{-x} b e^{-\frac{x}{b}}) \, dx
\]

\[
= 2b \int_{0}^{\infty} e^{-x} \, dx = \frac{1}{b} \ln b
\]

\[
= \frac{\ln 2}{b} \times \frac{1}{b} = \frac{1}{b} \cdot (\text{Ans})
\]
**Problem 3.** Let $X$ be an absolutely continuous random variable with the distribution function $F(x)$ and pdf $f(x)$.

Let $\alpha > 0$. Let $\alpha(x) = \frac{f(x)}{1 - F(x)}$, $\alpha > 0$.

$\alpha(x) = \text{Failure rate function or Hazard function}.$

Show that, $\alpha(x) = \text{constant}$ iff $X$ follows exponential dist.

**Soln:**

If part: Let $X$ is an exponential random variable.

i.e. $X \sim \text{E}(0, \theta)$

\[ f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x \geq 0 \]

\[ 1 - F(x) = e^{-x/\theta} \]

\[ \alpha(x) = \frac{f(x)}{1 - F(x)} = \frac{1}{\theta} \cdot \frac{e^{-x/\theta}}{e^{-x/\theta}} = \frac{1}{\theta} \]

Only if part: Suppose $\alpha(x) = \alpha$ (constant)

\[ \frac{f(x)}{1 - F(x)} = \alpha \]

on $\frac{\frac{d}{dx}F(x)}{1 - F(x)} = \alpha$

\[ \ln(1 - F(x)) = \alpha \ln(1 - F(x)) = -e^{\alpha x} + c \]

\[ 1 - F(x) = ke^{-\alpha x} \quad \text{[where, } k = e^c = \text{constant}] \]

\[ F(x) = 1 - ke^{-\alpha x} \]

\[ f(x) = k \alpha e^{-\alpha x} \]

Now, $F(0) = 1 - k$

\[ k = 1, \quad \alpha = 1 \]

\[ f(x) = \alpha e^{-x/\theta}, \quad \text{i.e. } f(x) = \frac{1}{\theta} e^{-x/\theta} \quad \text{[:: } \theta = \frac{1}{\alpha} \]
NORMAL DISTRIBUTION:

- Derivation of the p.d.f.: =>

The frequency curve obtained by approximating the histogram, takes the form like this:

\[
\begin{align*}
\frac{f(x)}{f(x)} & = \text{Frequency density} \\
x \rightarrow
\end{align*}
\]

In several real life situations, the density curve is symmetric bell shaped (WLG: let it be symmetric about 0').

1. The density curve is symmetric bell shaped (WLG: let it be symmetric about 0').
2. The density function has points of inflection and asymptotes at both the tails.

Suppose \( x \) is a continuous random variable having pdf \( f(x) = e^{-ax^2} \), \( a > 0 \). In order to reflect faithfully the above features,

Suppose, \( f(x) = e^{-ax^2} \)

\[
\begin{align*}
\int_{-\infty}^{\infty} f(x) dx &= 1 \\
\Rightarrow 2a & \int_{0}^{\infty} e^{-ax^2} dx = 1 \\
\Rightarrow 2a & \int_{0}^{\infty} e^{-ay^2} dy = 1 \\
\Rightarrow 2a & \int_{0}^{\infty} y e^{-ay^2} dy = 1 \\
\Rightarrow c &= \sqrt{\frac{\pi}{2a}} \\
\Rightarrow r &= \frac{\sqrt{a}}{\pi} \\
\Rightarrow \text{E}(x) &= 0 \quad \text{[as if it exists]}
\end{align*}
\]

\[
\begin{align*}
\text{Var}(x) &= \text{E}(x^2) - \text{E}(x)^2 \\
&= \int_{-\infty}^{\infty} x^2 e^{-ax^2} dx - 0 \\
&= 2a \int_{0}^{\infty} y^2 e^{-ay^2} dy \\
&= \frac{3\sqrt{\pi}}{2a^{3/2}} \\
&= \frac{3\sqrt{\pi}}{2a^{3/2}} \\
&= \sqrt{\frac{\pi}{2a}} \\
&= \sqrt{\frac{\pi}{2a}} \\
\Rightarrow \text{Var}(x) &= \frac{1}{2a}
\end{align*}
\]

\[
\begin{align*}
\Rightarrow a &= \frac{1}{2\sqrt{\pi}}, \quad c &= \frac{1}{\sqrt{2\pi}} \\
\Rightarrow \text{Var}(x) &= \frac{1}{2a}
\end{align*}
\]
\[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \]

Now, \( x' = x + \mu \); \( x' \) is a random variable having pdf \( f' \).

Pf of \( x' \) will be,

\[ P[x' = x] = f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \]

This is the p.d.f. of the Normal distribution with mean \( \mu \) and variance \( \sigma^2 \).

We denote the distribution by \( x \sim N(\mu, \sigma^2) \).

- **Definition:** An absolutely continuous random variable \( x \) defined over \( (-\infty, \infty) \) is said to follow normal distribution with parameters \( \mu, \sigma^2 \) if its pdf is given by,

\[ f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \]

We write \( x \sim N(\mu, \sigma^2) \).

**Problem:** If \( x \sim N(\mu, \sigma^2) \), then show that the distribution of \( x \) is symmetric about \( \mu \).

**Proof:**

\[ f(\mu + x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x+\mu-\mu}{\sigma} \right)^2} = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} \left( \frac{x}{\sigma} \right)^2} \]

The distribution is symmetric about \( \mu \).

\[ \therefore \mu + \frac{\sigma}{\sqrt{2}} = \mu \]

**Moments:**

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

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

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

\[ = 0 \quad \text{[\because it is an odd function]} \]

\[ \therefore E(x) = \mu \]

The \((2n+1)^{th}\) order central moment of \( x \) is

\[ \mu_{2n+1} = E((x-\mu)^{2n+1}) = 0 \]

All odd ordered central moments of normal distribution are zero."
\[ \mu_{2n} = E(x - \mu)^{2n} = \int_{-\infty}^{\infty} (x - \mu)^{2n} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - \mu)^2}{2}} dx \]

\[ = \int_{-\infty}^{\infty} x^{2n} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \]

Using the substitution \( u = \frac{x - \mu}{\sqrt{2}} \), we get

\[ \frac{1}{\sqrt{2}} \int_{-\infty}^{\infty} u^{2n} e^{-u^2} du \]

\[ = \frac{\sqrt{\pi}}{2^{n+1}} \Gamma(n + \frac{1}{2}) \]

For \( n = 1, \mu_2 = \mu^2 \)

For \( n = 2, \mu_4 = 3\mu^4 \)

Skewness \( \gamma_1 = 0 \)

Kurtosis \( \gamma_2 = \frac{3\mu^4}{\mu^2} - 3 = 0 \)

Normal distribution is perfectly skewed and mesokurtic.

\[ X \sim N(\mu, \sigma^2) \]

\[ E(x - \mu)^{2n-1} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (x - \mu)^{2n-1} e^{-\frac{(x - \mu)^2}{2}} dx \]

\[ = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (2\sigma)^{2n-1} e^{-\frac{(x - \mu)^2}{2}} dx \]

\[ = \frac{(2\sigma)^{2n-1}}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x - \mu)^2}{2}} dx \]

\[ = 0, \text{ since the integral } e^{2n-1} e^{-x^2/2} \text{ is odd.} \]

All odd central moment vanishes.
Moment Generating Function (MGF):

\[ MGF = M_C(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f(x) dx \]

Let, \( X \sim N(\mu, \sigma^2) \)

\[ e^{\mu t + \frac{1}{2}(\sigma^2 t^2)} \]

It is the pdf of \( N(\mu t, \sigma^2 t^2) \).

Central moment generating function:

\[ M_x(t) = \exp(\mu t + \frac{1}{2} \sigma^2 t^2) \]

\[ M(t) = 1 + \mu t + \frac{\mu^2 t^2}{2} + \cdots \]

Cumulant Generating Function:

\[ CGF = k(t) = \ln M(t) = \ln \left[ e^{\mu t + \frac{1}{2} \sigma^2 t^2} \right] \]

\[ k(t) = \mu t + \frac{\sigma^2 t^2}{2} \]

\[ k_1 = \mu \]

\[ k_2 = \sigma^2 \]

\[ k_3 = k_4 = \cdots = 0 \]
Recursion Relation for Central moments:

\[ \mu_{2n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu)^{2n} \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx \]

Differentiating both sides, we get,

\[ \frac{d}{dt} \mu_{2n} = \int_{-\infty}^{\infty} (x-\mu)^{2n-1} \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \cdot \left( -\frac{1}{\sigma^2} (x-\mu) \right) \, dx \]

\[ + \frac{1}{\sigma^2} \int_{-\infty}^{\infty} (x-\mu)^{2n} \cdot \left[ \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \cdot \left( -\frac{1}{\sigma^2} (x-\mu) \right) \right] \, dx \]

\[ = -\frac{1}{\sigma^2} \mu_{2n} + \frac{1}{\sigma^3} \mu_{2n+2} \]

\[ \Rightarrow \mu_{2n+2} = \int_{-\infty}^{\infty} \left[ \frac{d\mu_{2n}}{dt} + \frac{1}{\sigma^3} \mu_{2n} \right] \, dx \]

AH method:

\[ \mu_{2n} = \frac{(2n)!}{2^n} \frac{1}{(2n-1) \sigma^{2n}} \]

Thus,

\[ \mu_2 = \sigma^2 \]

\[ \mu_4 = 3 \sigma^4 \]

\[ \mu_6 = 15 \sigma^6 \]

\[ \mu_8 = 105 \sigma^8 \]

\[ \mu_{10} = 945 \sigma^{10} \]

\[ \mu_{12} = 10395 \sigma^{12} \]

Recursion Relation for Raw moments:

\[ \mu'_r = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^r \cdot e^{-\frac{1}{2} \left( \frac{x-\mu}{\sigma} \right)^2} \, dx \]

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

\[ = \frac{1}{\sigma^r} \mu_r + \int_{-\infty}^{\infty} x^r \cdot \frac{1}{\sqrt{2\pi}} \cdot \left( -\frac{1}{\sigma^2} (x-\mu) \right) \, dx \]

\[ = \frac{1}{\sigma^r} \mu_r + \frac{1}{\sigma^3} \mu_{r+2} - \frac{2r}{\sigma^3} \mu_{r+1} + \frac{r^2 \mu_r}{\sigma^3} \]
Mean Deviation about mean:

\[ MD_\mu (x) = E|X-\mu| \]
\[ = \int_{-\infty}^{\infty} \frac{|x-\mu|}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx \]
\[ = \frac{2\sigma}{\sqrt{2\pi}} \int_{0}^{\infty} z e^{-z^2/2} dz \]
\[ = \frac{2\sigma}{\sqrt{2\pi}} \sqrt{2\pi} = 2\sigma \]

Hence,

\[ \text{Variance} (|X-\mu|) = \text{E}[(X-\mu)^2] - [\text{E}(|X-\mu|)]^2 \]
\[ = \sigma^2 - \sigma^2 \frac{2}{\sqrt{2\pi}} \]
\[ = \sigma^2 (1 - \frac{2}{\sqrt{2\pi}}) \]
\[ \therefore \sqrt{\frac{\mu_2}{2}} = \text{S.D.} = \sigma \]
\[ \frac{MD}{\text{S.D.}} = \frac{2}{\sqrt{\pi}} = \frac{2}{1.77} < 1 \]

Problem 2: If \( X \sim N(\mu, \sigma^2) \) then show that:

\[ E[(X-\mu)g(x)] = \sigma \text{E}[g'(x)] \]

Hence, find the recursion relation for central moments.

**Solution:**

\[ E[(X-\mu)g(x)] = \int_{-\infty}^{\infty} (x-\mu)g(x) \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx \]
\[ = g(x) \times 0 - \int_{-\infty}^{\infty} g'(x) \left[ \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} dx \right] dx \]
\[ = - \int_{-\infty}^{\infty} g'(x) \left( \sigma^2 \left[ \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2}} \right] \right) dx \]
\[ = - \sigma^2 \text{E}[g'(x)] \quad \text{[Proved]} \]

Now, let, \( g(x) = (x-\mu)^{2n-1} \)
\[ g'(x) = (2n-1)(x-\mu)^{2n-2} \]
\[ E(x-\mu)^{2n} = \sigma^2 E[(2n-1)(x-\mu)^{2n-2}] \]
\[ \therefore E(x-\mu)^{2n} = \sigma^2 (2n-1) E(x-\mu)^{2n-2} \]
Mode & Point of Inflation:

\[
\begin{align*}
\hat{f}(x) &= \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}} \left(\frac{x-\mu}{\sigma}\right)^2, \\
\hat{f}'(x) &= \frac{1}{\sqrt{2\pi}} \cdot \left(-\frac{1}{2}\right) \left(\frac{x-\mu}{\sigma}\right) \cdot \frac{1}{\sigma^2} \cdot 2(x-\mu), \\
\hat{f}''(x) &= -\hat{f}(x) \left(\frac{x-\mu}{\sigma}\right), \\
&= - \left[ \left(\frac{x-\mu}{\sigma}\right) \hat{f}'(x) + \hat{f}(x) \cdot \frac{1}{\sigma} \right], \\
&= - \left[ -\left(\frac{x-\mu}{\sigma}\right)^2 \hat{f}(x) + \hat{f}(x) \cdot \frac{1}{\sigma} \right], \\
&= \frac{\hat{f}(x)}{\sigma^2} \left[ \left(\frac{x-\mu}{\sigma}\right)^2 - 1 \right].
\end{align*}
\]

Now, \( \hat{f}'(x) = 0 \Rightarrow -\hat{f}(x) \left(\frac{x-\mu}{\sigma}\right) = 0 \)
\[\Rightarrow \alpha = \mu.\]

\[\hat{f}''(\mu) = -\frac{\hat{f}(\mu)}{\sigma^2} < 0.\]
\[\therefore \hat{f}(x) \text{ has a maxima at } \alpha = \mu. \]

\[\therefore \text{Mode of the distribution } \alpha = \mu.\]

Now, \( \hat{f}''(x) = 0 \Rightarrow \frac{\hat{f}(x)}{\sigma^2} \left[ \left(\frac{x-\mu}{\sigma}\right)^2 - 1 \right] = 0 \)
\[\Rightarrow \left(\frac{x-\mu}{\sigma}\right)^2 = 1, \]
\[\Rightarrow \alpha = \mu \pm \sigma.\]

\[\therefore \hat{f}(x) \text{ has a point of inflection at } \alpha = \mu \pm \sigma.\]

Now, \( \hat{f}''(x) < 0 \) whenever \( \alpha \to +\infty \), \( \alpha \to -\infty \).
\[\therefore \sigma < 0 \Rightarrow \alpha \neq \mu. \hat{f}(x) \not\to 0 \text{ as } \alpha \to \pm\infty.\]

\[\therefore \hat{f}(x) \text{ has no asymptote at } \alpha = \pm\infty.\]

Example:

Let \( X \sim N(\mu, \sigma^2) \) then show that \( E|X-\alpha| > \sqrt{\frac{2}{\pi}} \) for any \( \alpha \).

Solution:

Mean deviation is minimum when it is measured about the median. Here median = \( \mu \).
\[\therefore E|X-\alpha| = E|X-M(x)| = E|X-\mu|, \]
\[\geq \sqrt{\frac{2}{\pi}} \text{ when } \alpha \sim N(\mu, \sigma^2).\]
Reproductive Property:

Let, \( X_1 \sim N(\mu_1, \sigma_1^2) \)
\( X_2 \sim N(\mu_2, \sigma_2^2) \)
\( X_1 \) and \( X_2 \) are independent random variables.

Define, \( S = l_1 X_1 + l_2 X_2 + l_0 \), where \( l_i \)'s are known constants.

\[
M_S(t) = E(e^{tS}) = E(e^{t(l_0 + l_1 X_1 + l_2 X_2)})
= E(e^{t l_0}) E(e^{t l_1 X_1}) E(e^{t l_2 X_2})
= e^{t l_0} M_{X_1}(t l_1) M_{X_2}(t l_2)
= e^{t l_0} e^{t l_1 \mu_1 + \frac{t^2}{2} \sigma_1^2} e^{t l_2 \mu_2 + \frac{t^2}{2} \sigma_2^2}
= e^{t l_0 + l_1 \mu_1 + l_2 \mu_2 + \frac{t^2}{2} \sigma_1^2 + \sigma_2^2}
= e^{t \mu_S + \frac{t^2}{2} \sigma_S^2}
\]

As MGF uniquely determines normal distribution.

\( S \sim N(\mu_S, \sigma_S^2) \)

i.e. \( S \sim N(l_0 + l_1 \mu_1 + l_2 \mu_2, l_1 \sigma_1^2 + l_2 \sigma_2^2) \)

In general, if \( X_i \sim N(\mu_i, \sigma_i^2), i=1(1)n \),
where \( X_i \)'s are independently distributed, then,

\( S_n = \sum_{i=1}^{n} X_i \sim N(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2) \)

If further, \( X_i \)'s are identically distributed with common mean \( \mu \) and common variance \( \sigma^2 \), then,

\( S_n \sim N(n \mu, n \sigma^2) \)

and \( \overline{X}_n = \frac{S_n}{n} \sim N(\mu, \frac{\sigma^2}{n}) \).

Some specific cases:

1. \( X \sim N(\mu, \sigma^2) \) \( \Rightarrow X + c \sim N(\mu + c, \sigma^2) \); \( c = \text{constant} \)
2. \( X \sim N(\mu, \sigma^2) \) \( \Rightarrow cX \sim N(c\mu, c^2 \sigma^2) \); \( c \neq 0 \)
   \( \Rightarrow \text{If } c = 0, \text{ then } P[X = 0] = 0 \)
   \( \Rightarrow X \) will be a degenerate random variable.
3. \( X \sim N(\mu, \sigma^2) \) \( \Rightarrow -X \sim N(-\mu, \sigma^2) \)
   \( \Rightarrow \text{if } \mu = 0, \overline{X} \equiv -X \) [identical distribution]
4. \( X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \)
   \( \Rightarrow X_2 - X_1 \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \)
   \( \Rightarrow \text{if } \mu = 0, X_1 + X_2 \overset{D}{=} X_1 - X_2 \)
• Standard Normal Variate:

$x$ is said to be a standard normal variate if

$x \sim N(0, 1)$.

$x \sim N(\mu, \sigma^2) \iff \frac{x - \mu}{\sigma} \sim N(0, 1)$.

Let, $z = \frac{x - \mu}{\sigma}$, then $z \sim N(0, 1)$

The distribution function of $z$ is $\Phi(z) = P[Z \leq z]$.

$\Gamma(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}} dt$.

$\frac{d}{dz} \Phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

$\therefore \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$.

Properties of $\phi(z)$ and $\Phi(z)$:

1. $\phi(z) = \phi(-z), \forall z$
2. $\Phi(z) = 1 - \Phi(-z), \forall z$

Proof:

$\Phi(z) = \int_{-\infty}^{z} \phi(u) du$

$\Phi(-z) = \int_{-\infty}^{-z} \phi(u) du$

$= \int_{-\infty}^{z} \phi(u) du$

$\therefore \Phi(z) + \Phi(-z) = \int_{-\infty}^{z} \phi(u) du + \int_{-\infty}^{z} \phi(u) du$

$= 1, \forall z$

Note, $\Phi(0) = \frac{1}{2}$. 

Implication: If $\Phi(z), z > 0$ is given, then we can find $\phi(z)$ from the above formula.
Notes -
\[ P[X_1 > X_2] = P[X_1 - X_2 > 0] \]
\[ = P \left[ \frac{(X_1 - X_2) - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2}} > \frac{\mu_1 - \mu_2}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right] \]
\[ = P \left[ Z > \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right] , \quad Z \sim N(0,1) \]
\[ = 1 - \Phi \left( \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right) \]

Now, if \( \mu_1 = \mu_2 \), then, \[ P[X_1 > X_2] = 1 - \Phi(0) \]

\[ \boxed{\star \text{ Problem 8.} \quad \exists X \sim N(0,1) \text{ and } E(\Phi(X)) \text{ and } V(\Phi(X))} \]

**Ans:** \[ E(\Phi(X)) = \int_{-\infty}^{\infty} \Phi(X) \phi(X) \, dx = \int_{-\infty}^{\infty} z \, dz \]
\[ = \frac{1}{2} \]
\[ E(\Phi(X)^2) = \int_{-\infty}^{\infty} (\Phi(X))^2 \phi(X) \, dx = \int_{-\infty}^{\infty} (\Phi(X))^2 \, dx \]
\[ = \frac{1}{3} \]
\[ V(\Phi(X)) = \frac{1}{2} - \frac{1}{3} = \frac{1}{12} \]

\[ \begin{align*}
Z &= \Phi(X) \sim R(0,1) \\
E(Z) &= \int_{-\infty}^{\infty} z \phi(z) \, dz = \frac{1}{2} \\
E(Z^2) &= \int_{-\infty}^{\infty} z^2 \phi(z) \, dz = \frac{1}{3} \\
V(Z) &= \frac{1}{2} - \frac{1}{3} = \frac{1}{12} \\
\end{align*} \]

\[ \star \text{ Problem 4.} \quad \exists X \sim N(0,1) \text{ and } E[X \Phi(X)] \text{ and } E[X^2 \Phi(X)] \]

**Ans:** \[ E(X \Phi(X)) = \int_{-\infty}^{\infty} x \Phi(X) \phi(x) \, dx = \int_{-\infty}^{\infty} \Phi(X) \phi(x) \, dx \]
\[ = \left[ \Phi(X) - \int \Phi(X) \, dx \right] \int_{-\infty}^{\infty} \phi(x) \, dx + \int \phi(x) \, dx \]
\[ = 0 + \int \frac{1}{2\pi} e^{-\frac{x^2}{2}} \, dx \]
\[ = \frac{1}{2\sqrt{\pi}} \]
Problems. Let $Z \sim N(0,1)$ and $G_i(z) = P[0 \leq Z \leq z]$. Then show that:

\[
\Phi(z) \left(1 - \Phi(z)\right) = \frac{1}{4} - G_i(z).
\]

\[
\text{Solved} \rightarrow G_i(z) = \Phi(z) - \Phi(0) = \Phi(z) - \frac{1}{2}
\]

\[
\Phi(z) = G_i(z) + \frac{1}{2}.
\]

\[
\Phi(z) \left[1 - \Phi(z)\right] = \left[G_i(z) + \frac{1}{2}\right] \left[\frac{1}{2} - G_i(z)\right].
\]

Problems. If $X \sim N(\mu, 1)$, $\mu > 0$, then show that:

\[
E\left[\frac{1 - \Phi(x)}{\phi(x)}\right] = \frac{1}{\mu},
\]

so hence $\phi$, $\Phi$, respectively, denote the cdf and pdf of $N(\mu, \sigma^2)$.
Problem 7. If $X \sim N(0, \sigma^2)$ then show that:

$$P[X > x + \frac{c}{\sigma}] < e^{-c^2/2\sigma^2}.$$ 

**Solution**:

$$P[X > x + \frac{c}{\sigma}] = \int_{x + \frac{c}{\sigma}}^{\infty} \frac{1}{\sqrt{2\pi} \sigma} e^{-t^2/2\sigma^2} dt$$

Let us transform $z = t - \frac{c}{\sigma}$

Now,

$$(z + \frac{c}{\sigma})^2 = z^2 + 2zc \frac{c}{\sigma} + \frac{c^2}{\sigma^2} > z^2 + 2zc \frac{c}{\sigma}$$

$$e^{z^2 + 2zc \frac{c}{\sigma}} > e^{z^2 + 2zc}$$

$$e^{-\frac{c^2}{2\sigma^2}} < e^{-2zc}$$

$$\therefore \frac{P[X > x + \frac{c}{\sigma}]}{P[X > x]} < e^{-c/\sigma}.$$ 

2. QUARTILE DEVIATION OF NORMAL DISTRIBUTION:

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

$$P[X \leq \bar{q}_{1/2}] = \frac{1}{2}$$

Now, $P[X \leq \bar{q}_p] = P$

$$P[X \leq \bar{q}_{3/4}] = \frac{3}{4}$$

Note, $\Phi(\infty) = 1 \Rightarrow 1 - \Phi(\infty) = \frac{1}{2}$

$$\Rightarrow \Phi(-\infty) = \frac{3}{4}$$

$$\Rightarrow -\infty = \Phi^{-1}(\frac{3}{4})$$

$$\Rightarrow \infty = -\Phi^{-1}(\frac{3}{4})$$

$$\therefore \bar{q}_{1/2} = \mu + \sigma \Phi^{-1}(\frac{1}{2})$$

and $\bar{q}_{3/4} = \mu + \sigma \Phi^{-1}(\frac{3}{4}).$
**Lognormal Distribution**

In order to graduate (smoothing out the irregularities) an observed income distribution is log-normal (and possibly tailed) distribution of an adopted. For a typical results distribution usually a mixture of the two employed to the hump (mode) of the distribution usually a lognormal density is fitted and beyond it a Pareto density is used.

- **Lognormal Distribution:**

  $x$ is said to have a lognormal distribution if $\ln x$ is normally distributed.

  From the picture it is clear that the lognormal distribution is a positively skewed and bell-shaped distribution.

  **Definition:** An absolutely continuous random variable $X$ defined over $(0, \infty)$ is said to follow lognormal distribution with parameter $\mu$ and $\sigma$ if its pdf is given by:

  $$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln x - \mu}{\sigma}\right)^2\right), \quad 0 < x < \infty$$

  $$\sigma > 0$$

  We write $X \sim \mathcal{L}(\mu, \sigma^2)$.

  **Note:** $\ln X \sim \mathcal{N}(\mu, \sigma^2)$.

  $$\Pr [x \leq \alpha] = \Pr [\ln x \leq \ln \alpha] \quad [\text{since } \ln x \text{ is an increasing function of } x]\)$$

  $$= \int_0^{\ln \alpha} \phi \left(\frac{\ln x - \mu}{\sigma}\right) dx \quad \text{[PDF of } \mathcal{N}(\mu, \sigma^2)]$$

  **Probability Density:**

  $$f(x) = \frac{1}{\sigma x \sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{\ln x - \mu}{\sigma}\right)^2\right), \quad 0 < x < \infty$$
Result:
\[ x \sim \Lambda(\mu, \sigma^2) \]
\[ \iff \quad \ln x \sim N(\mu, \sigma^2) \]

Proof:
- **If part:** \( \iff \) Let \( Y = \ln x \sim N(\mu, \sigma^2) \).

To show, \( x \sim \Lambda(\mu, \sigma^2) \).

Distribution function of \( x \) is

\[ F(x) = P[x \leq x] = P[\ln x \leq \ln x] = G(\ln x), \text{ where } G \text{ denotes the DF of } Y. \]

\[ \therefore \quad \text{The pdf of } x \text{ is,} \]

\[ f(x) = \frac{d}{dx} G(\ln x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}} \quad ; \quad 0 < x \]

Only if part:

Let \( x \sim \Lambda(\mu, \sigma^2) \).

To show, \( Y = \ln x \sim N(\mu, \sigma^2) \).

The distribution function of \( Y \) is,

\[ G_Y(y) = P[Y \leq y] = P[\ln x \leq y] = P[x \leq e^y] \]

\[ \therefore \quad F \text{ is distribution function of } x. \]

The pdf of \( Y \) is,

\[ g(y) = \frac{d}{dy} F(e^y) = e^y f(e^y) \]

\[ = e^y \cdot \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y - \mu)^2}{2\sigma^2}} \quad ; \quad -\infty < y < \infty \]

\[ \therefore \quad g(y) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y - \mu)^2}{2\sigma^2}} \quad ; \quad -\infty < y < \infty \]

Moments:

Let \( x \sim \Lambda(\mu, \sigma^2) \).

Then the rate order raw moment of \( x \) is,

\[ \mu_r = E(x^n) = E(e^{ry}) = M_Y(n) \quad ; \quad Y = \ln x \]

\[ = e^{\mu + \frac{1}{2} \sigma^2} \quad ; \quad x = e^y \]

Mean = \( e^{\mu + \frac{1}{2} \sigma^2} = e^{\mu_{\ln} + \frac{1}{2} \sigma^2} = [\text{where } \delta = e^{\frac{\sigma^2}{2}}] \)
Mode of the distribution:

\[ f(x) = \frac{1}{\alpha \sqrt{2\pi}} \cdot e^{-\frac{1}{2} \left( \frac{(\ln x - \mu)^2}{\sigma^2} \right)} \]

\[ f'(x) = \frac{1}{\alpha \sqrt{2\pi}} \cdot \left[ -\frac{1}{2} \left( \frac{(\ln x - \mu)^2}{\sigma^2} \right) \cdot 2 \cdot (-\frac{1}{x}) \cdot \frac{(\ln x - \mu)}{\sigma^2} \right] \cdot \frac{1}{\alpha} \cdot e^{-\frac{1}{2} \left( \frac{(\ln x - \mu)^2}{\sigma^2} \right)} \]

\[ \Rightarrow \left[ -\frac{1}{x} \cdot \ln x - \frac{\mu}{\sigma^2} \right] = 0 \quad \text{[\because x > 0]} \]

\[ \Rightarrow \ln x - \frac{\mu}{\sigma^2} = 0 \]

\[ \Rightarrow x = e^{\mu - \sigma^2} \]

Now, \( f''(x) = -f(x) \cdot \left( \frac{\ln x - \mu}{\sigma^2} + 1 \right) \)

\[ \Rightarrow f''(e^{\mu - \sigma^2}) = f(e^{\mu - \sigma^2}) \cdot \left( \frac{\ln e^{\mu - \sigma^2} - \mu}{\sigma^2} + 1 \right) = f(e^{\mu - \sigma^2}) \cdot \left( \frac{\mu - \mu}{\sigma^2} + 1 \right) = f(e^{\mu - \sigma^2}) \cdot \left( \frac{0}{\sigma^2} + 1 \right) = f(e^{\mu - \sigma^2}) \cdot \left( 1 \right) = f(e^{\mu - \sigma^2}) > 0 \]

\[ \therefore f(x) \text{ has a maxima at } x = e^{\mu - \sigma^2} \]

\[ \therefore \text{ Mode } = e^{\mu - \sigma^2} \]

Reproductive Property of Lognormal Distribution:

Suppose \( x_1 \) and \( x_2 \) are independently distributed lognormal variables, where:

\[ x_1 \sim \Lambda \left( \mu_1, \sigma_1^2 \right) \]

\[ x_2 \sim \Lambda \left( \mu_2, \sigma_2^2 \right) \]

Define, \( Y = \alpha x_1^b x_2^c \), \( a > 0 \)

\( a, b, c \) are non-zero stochastic quantities.

Clearly, \( \ln Y = \ln a + b \ln x_1 + c \ln x_2 \)

\[ \therefore \ln x_1 \text{ and } \ln x_2 \text{ are independent } \sim N \left( \mu_1, \sigma_1^2 \right), \text{ i.e., } z_1, z_2 \]

Then the reproductive property of normal distribution, \( \ln Y \sim N \left( \ln a + b \mu_1 + c \mu_2, b \sigma_1^2 + c \sigma_2^2 \right) \)

\[ \therefore Y \sim \Lambda \left( \ln a + b \mu_1 + c \mu_2, b \sigma_1^2 + c \sigma_2^2 \right) \]

\[ \therefore x_1, x_2 \text{ are independent } \sim \Lambda \left( \mu_1, \sigma_1^2 \right), \text{ i.e., } z_1, z_2 \]
Some particular cases:

1) \( X \sim \Lambda (\mu, \sigma^2) \)
   \( aX \sim \Lambda (a\mu, a\sigma^2) \), \(-a > 0\)

2) \( X \sim \Lambda (b\mu, b\sigma^2) \)
   \( \frac{1}{X} \sim \Lambda (-\mu, \sigma^2) \).
   \( X = \frac{1}{X} \) [if \( \mu = 0 \)]

3) \( X_1, X_2 \sim \Lambda (\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \)
   \( \frac{X_1}{X_2} \sim \Lambda (\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2) \)
   \( X_1X_2 \sim \frac{x_1}{x_2} \) [if \( \mu_2 = 0 \)]

**Problem 1.** If \( X \sim \Lambda (\mu, \sigma^2) \) then \( AM > GM > HM \).

*Solution.*

\[ AM = E(X), \quad GM = e^{E(\ln X)}, \quad HM = \frac{1}{E\left(\frac{1}{X}\right)} \]

Hence, \( f(x) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{1}{2} \left( \frac{x - \mu}{\sigma} \right)^2} \)

\[ E(X) = e^{\mu + \frac{\sigma^2}{2}} \]

\[ E(\ln X) = \int_{-\infty}^{\infty} \ln x \cdot f(x) \, dx \]

\[ = e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} \frac{e^{\frac{x^2}{2}}}{\sqrt{2\pi}} \, dx \]

\[ = e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{e^{\frac{x^2}{2}}}{\sqrt{2\pi}} \, dx \]

\[ = e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{e^{z^2}}{\sqrt{2\pi}} \, dz \]

\[ = e^{-\frac{\sigma^2}{2}} \]

\[ E\left(\frac{1}{X}\right) = \int_{-\infty}^{\infty} \frac{1}{x} \cdot f(x) \, dx \]

\[ = e^{-\frac{\sigma^2}{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} \, dx \]

\[ = e^{-\frac{\sigma^2}{2}} \]

\[ GM = e^{\mu} \]

\[ HM = e^{\mu - \frac{\sigma^2}{2}} \]

\[ AM > GM > HM \]
**Cauchy Distribution:**

**Derivation of the p.d.f.:**

Hence, \( \Theta \sim \text{Uniform} \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \).

Clearly, \( x = \tan \Theta \).

\[
\frac{d}{dx} P[\Theta \leq \tan^{-1} x] = \frac{1}{1 + x^2} \\
= \frac{1}{\pi} \frac{1}{1 + x^2} \quad \text{for} \quad x \in \mathbb{R}
\]

\( \Rightarrow f(x) = \frac{1}{\pi} \frac{1}{1 + x^2} \quad \text{for} \quad x \in \mathbb{R} \) \( \Rightarrow x \sim \text{Cauchy} \).

The above is the pdf of the **Standard Cauchy Distribution**.

Clearly, Cauchy distribution is symmetrical about zero, i.e., median of the distribution is zero, moreover the distribution is bell-shaped, i.e., mode is also zero but the mean of the distribution does not exist.

- **Cauchy Density Function:**

The tails of the Cauchy distribution are heavy, i.e., the tail probabilities are significant.

Here it can be easily verified that \( x \sim \text{Standard Cauchy} \iff \frac{1}{x} \sim \text{Standard Cauchy} \).

It is evident that if \( x \sim \text{the standard Cauchy distribution} \), then distribution of \( x \) and \( -x \) will be identical (**Insular feature**).
**ONE PARAMETER CAUCHY DISTRIBUTION:**

A random variable $X$ is said to be Cauchy variable of median $\mu$ if the pdf of $X$ is of the form

$$f(x) = \frac{1}{\pi} \left( \frac{1}{1 + (x - \mu)^2} \right)$$

Clearly, the distribution is symmetric about $\mu$.

We denote, $X \sim \text{Cauchy with median} \mu$.

$\Rightarrow (X - \mu) \sim \text{standard Cauchy}$.

**TWO PARAMETER CAUCHY DISTRIBUTION:**

An absolutely continuous random variable $X$ defined over $(-\infty, \infty)$ is said to follow Cauchy distribution with parameters $\mu$ and $\gamma$.

- Median = $\mu$,
- Scale = $\gamma$

if its pdf is given by,

$$f(x) = \frac{1}{\pi \gamma} \left( \frac{1}{1 + (x - \mu)^2 / \gamma^2} \right); \quad -\infty < x < \infty,$$

$$-\infty < \mu < \infty,$$

$$\gamma > 0.$$

We denote, $X \sim c(\mu, \gamma)$.

$$\Rightarrow (X - \mu) \sim c(0,1).$$

**Result:** If $X \sim c(\mu, \gamma)$, then show that $\frac{X - \mu}{\gamma} \sim c(0,1)$.

**Solution:**

$$X \sim c(\mu, \gamma)$$

$$f_X(x) = \frac{1}{\pi \gamma} \left( \frac{1}{1 + (x - \mu)^2 / \gamma^2} \right), \quad \alpha \in \mathbb{R}, \quad \mu \in \mathbb{R}, \quad \gamma > 0.$$

$$F_Z(z) = P[Z \leq z] = P \left[ \frac{X - \mu}{\gamma} \leq z \right] = P[X \leq \mu + \gamma z] = F_X(\mu + \gamma z)$$

$$\Rightarrow f_Z(z) = \frac{d}{dz} F_Z(z) = \frac{d}{dz} F_X(\mu + \gamma z) = \gamma \cdot f_X(\mu + \gamma z)$$

$$= \frac{1}{\pi \gamma} \left( \frac{1}{1 + ((x - \mu) / \gamma)^2} \right); \quad -\infty < x < \infty,$$

$$-\infty < \mu < \infty,$$

$$0 < \gamma.$$
Reproductive Property of Cauchy Distribution:

Suppose \( X_1, X_2, \ldots, X_n \) are independently distributed Cauchy variables where, \( X_i \sim \text{C}(\mu_i, \gamma_i) \) \( \forall i = 1, 2, \ldots, n \).

Define, \( S_n = \sum_{i=1}^{n} X_i \),

then \( S_n \sim \left( \sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \gamma_i \right) \).

In particular, if \( X_i \)'s are i.i.d. Cauchy variables with median \( \mu \) and scale \( \gamma \) then,

\[
S_n \sim \text{C}(n\mu, n\gamma) \iff \bar{X}_n = \frac{S_n}{n} \sim \text{C}(\mu, \gamma).
\]

i.e. in case of sampling from Cauchy \((\mu, \gamma)\) distribution the sampling distribution of the sample mean will be same as the parent distribution.

Note:— Like normal distribution, Cauchy distribution is also a stable distribution. (Proof by Central Limit Theorem).

Proof of the Reproductive property:

Result:— If \( X \sim \text{C}(0,1) \) and \( Y \sim \text{C}(0,1) \) then \( Z = X + Y \sim \text{C}(0,2) \).

Proof:— \( X \sim \text{C}(0,1) \) \( \Rightarrow f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2} \); \( -\infty < x < \infty \)

\( Y \sim \text{C}(0,1) \) \( \Rightarrow f(y) = \frac{1}{\pi} \cdot \frac{1}{1+y^2} \); \( -\infty < y < \infty \)

\[
Z = X + Y, \quad \int_{-\infty}^{\infty} \frac{1}{2\pi} \cdot \frac{1}{1+x^2} \cdot \frac{1}{1+(z-x)^2} \, dx
\]

Now,

\[
\frac{1}{(1+x^2)(1+(z-x)^2)} = \frac{1}{2\pi^2} \left[ \frac{2x}{1+x^2} + \frac{2z-2x}{1+(z-x)^2} \right]
\]

so that,

\[
f_Z(z) = \frac{1}{2\pi} \cdot \frac{1}{2\pi^2} \left[ \frac{2 \log \left( \frac{1+x^2}{1+(z-x)^2} \right) + 2 \tan^{-1} x + 2 \tan^{-1} (z-x)}{1+(z-x)^2} \right]
\]

\[
= \frac{1}{\pi} \cdot \frac{2}{(z+2)^2}, \quad -\infty < z < \infty
\]

\[\therefore X + Y \sim \text{C}(0,2). \quad \text{[Proved]} \]
Some features of two-parameter Cauchy Distribution

**Quartiles**

The distribution function of $X$ is,

$$F(x) = \int_{-\infty}^{x} f(t) \, dt = \int_{-\infty}^{x} \frac{1}{\pi \sigma} \frac{x - \mu}{(x - \mu)^2 + \sigma^2} \, dt$$

$$= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dt}{1 + \left(\frac{t - \mu}{\sigma}\right)^2}$$

$$= \frac{1}{\pi} \left[ \tan^{-1}\left(\frac{x - \mu}{\sigma}\right) + \frac{\pi}{2} \right]$$

$$= \frac{1}{\pi} \tan^{-1}\left(\frac{x - \mu}{\sigma}\right) + \frac{1}{2}$$

$$F\left(\frac{\alpha}{4}\right) = \frac{1}{\pi} \tan^{-1}\left(\frac{\alpha}{4}\right) + \frac{1}{2} = P$$

$$\Rightarrow \tan^{-1}\left(\frac{\alpha}{4}\right) = \pi\left(P - \frac{1}{2}\right)$$

$$\Rightarrow \frac{\alpha}{4} = \pi\left(P - \frac{1}{2}\right)$$

$$\Rightarrow \alpha = 4\pi\left(P - \frac{1}{2}\right)$$

- Quartile Deviation $= \frac{\alpha}{2} = \pi\left(P - \frac{1}{2}\right)$

- Quartile $Q_1 = \mu - \frac{\pi}{2}$

- Quartile $Q_2 = \mu$

- Quartile $Q_3 = \mu + \frac{\pi}{2}$

**Mode**

$$f(x) = \frac{1}{\pi} \frac{1}{x^2 + (x - \mu)^2}$$

$$f'(x) = \frac{1}{\pi} \frac{-2(x - \mu)}{(x^2 + (x - \mu)^2)^2} = 0$$

$$\Rightarrow x = \mu$$

$$f''(x) = \frac{1}{\pi} \frac{(-2)(x^2 + (x - \mu)^2) - 2(x - \mu)^2}{(x^2 + (x - \mu)^2)^3}$$

$$f''(\mu) = \frac{1}{\pi} \frac{-2\sigma^2}{4d^4} < 0$$

$$\Rightarrow f(x) \text{ has its maxima at } x = \mu$$

$$\therefore \text{ Mode } = \mu$$
Maximum ordinate: at $x = \mu$,

\[ f(\mu) = \frac{1}{\pi \sigma}. \]

**Problem 1.** If $X \sim \mathcal{N}(\mu, \sigma)$, then show that the distribution is symmetric about $\mu$.

**Ans:**

\[ f(\mu + h) = \frac{e^{-\frac{(h+\mu)^2}{2\sigma^2}}}{\pi \sigma} = f(\mu - h), \]

i.e., the distribution is symmetric about $\mu$.

**Problem 2.** If $X \sim \mathcal{N}(0,1)$ then show that, $E(X^n)$ exists if $|n| < 1$.

**Proof:**

\[ E[X^n] = \int_{-\infty}^{\infty} x^n \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx \]

For a causal distribution,

\[ = \int_{0}^{\infty} x^n \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = \left[ \text{assuming } \int_{0}^{\infty} \frac{x^n}{\sqrt{2\pi}} dx \text{ converges} \right] \]

For a causal distribution,

\[ = \frac{1}{\pi} \int_{0}^{\infty} x^n e^{-\frac{x^2}{2}} dx \text{ converges } \]

Let, $\alpha = 2$, $2x dx = d\alpha$.

\[ = \frac{1}{\pi} \int_{0}^{\infty} x^{\frac{n}{2} - 1} e^{-\frac{x^2}{2}} dx \]

\[ = \frac{1}{\pi} \int_{0}^{\infty} \frac{2^{\frac{n}{2} - 1}}{\left(\frac{1}{2} + \frac{n}{2}\right)} x^{\frac{n}{2} - 1} e^{-\frac{x^2}{2}} dx \]

\[ = \left\{ \begin{array}{ll}
\frac{1}{\pi} & \text{if } \frac{n}{2} - 1 < \frac{1}{2} \Rightarrow 0 < n < 1
\end{array} \right. \]

\[ \therefore E[X^n] \text{ exists for all } n \text{ if } |n| < 1. \]

**Note:** If $X \sim \mathcal{N}(0,1)$ then $E[X^n]$ exists for $|n| < 1$.

**Problem 3.** If $X \sim \mathcal{N}(0,1)$ then show that $\frac{1}{X} \sim \mathcal{N}(0,1)$. 

**Proof:** Let $Y = \frac{1}{X}$, $F_Y(y) = P\left[ Y \leq y \right]

\[ = P\left[ X \leq \frac{1}{y} \right] \]

\[ = 1 - P\left[ X > \frac{1}{y} \right] \]

\[ = 1 - \int_{\frac{1}{y}}^{\infty} \frac{1}{\sqrt{2\pi}} \int \left[ 1 - \frac{1}{\sqrt{2\pi}} \right] \left[ \int \right] \]

\[ = 1 - \left\{ \begin{array}{ll}
\frac{1}{\pi} & \text{if } \frac{1}{y} - 1 \leq \frac{1}{2} \Rightarrow 0 < y < 1
\end{array} \right. \]

\[ \therefore Y \sim \mathcal{N}(0,1). \]
Problem 1. If \( x \sim \text{e}(0,1) \) then show that the MGF of \( x \) does not exist.

**Proof:**

\[
\text{MGF} = M_X(t) = E(e^{tx}) = \int_{-\infty}^{\infty} e^{tx} f_X(x) dx
\]

\[
= \int_{-\infty}^{0} e^{tx} \frac{1}{\pi(1+x^2)} dx + \int_{0}^{\infty} e^{tx} \frac{1}{\pi(1+x^2)} dx
\]

\[
= \int_{-\infty}^{0} \frac{e^{tx}}{\pi(1+x^2)} dx + \int_{0}^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx
\]

\[
= I_1 + I_2
\]

\[
I_1 = \int_{-\infty}^{0} \frac{e^{tx}}{\pi(1+x^2)} dx < \int_{-\infty}^{0} \frac{1}{\pi(1+x^2)} dx \quad \text{[for } t > 0, e^{tx} < 1 \text{ as } x \text{ is negative]}
\]

\[
= \frac{1}{2}
\]

\[
I_1 \text{ exists if } t > 0.
\]

\[
I_2 = \int_{0}^{\infty} \frac{e^{tx}}{\pi(1+x^2)} dx \quad \text{[if } t < 0 \text{, then } e^{tx} > tx]
\]

\[
= \frac{t}{2\pi} \int_{0}^{\infty} \frac{2x}{1+x^2} dx
\]

\[
= \frac{t}{2\pi} \lim_{s \to \infty} \left[ \log(1+x^2) \right]_0^s
\]

\[
= \frac{t}{2}\log(1+t^2)
\]

\[
I_2 \text{ does not exist for } t < 0.
\]

\[
I_2 \text{ exists when } t > 0, I_2 \text{ does not exist when } t < 0.
\]

Hence, there does not exist any \( t \) such that \( |t| < b \) for which \( E(e^{tx}) \) exists.

Problem 5. A straight line \( AB \) is free to move through a fixed point \( A \) with co-ordinates \((0, \mu)\) and the length of the intersection \( x \) it makes with the \( x \)-axis is noted. \( AB \) makes an angle \( \theta \) with the \( y \)-axis assuming that \( \theta \) has an uniform distribution between \(-\frac{\pi}{2}\) to \( \frac{\pi}{2}\). Find the distribution of \( x \).

Solution:

\[
\frac{x}{\mu} = \tan \theta
\]

\[
\Rightarrow x = \mu \tan \theta
\]
The distribution function of $X$ is,

$$F_x(x) = P \left[ X \leq x \right]$$

$$= P \left[ \tan^{-1} \frac{z}{\mu} \leq x \right]$$

$$= P \left[ \theta \leq \tan^{-1} \frac{x}{\mu} \right]$$

$$= \int_{-\pi/2}^{\tan^{-1} \frac{x}{\mu}} q(\theta) d\theta$$

$$= \int_{-\pi/2}^{\tan^{-1} \frac{x}{\mu}} \frac{1}{\pi} d\theta$$

$$= \left. \frac{\theta}{\pi} \right|_{-\pi/2}^{\tan^{-1} \frac{x}{\mu}}$$

$$= \frac{\tan^{-1} \frac{x}{\mu}}{\pi} + \frac{1}{2}$$

\[ \therefore \text{pdf of } X \text{ is, } \frac{d(\theta)}{\pi} = \frac{1}{\mu} \left( 1 + \left( \frac{x}{\mu} \right)^2 \right)^{-1} \]

\[ X \sim \text{C}(0, \frac{1}{\mu}) \]

Problems. If $X \sim \text{C}(0, \frac{1}{\mu})$, $Y = \frac{2x}{1+x^2}$, find $E(Y)$ and $V(Y)$.

\[ \text{Ans.} \]

$$f(x) = \frac{1}{\pi} (1+\alpha x)^{-2} I_x(-\infty, \infty)$$

$$E(Y) = E \left( \frac{2x}{1+x^2} \right) = \int_{-\infty}^{\infty} \frac{2x}{1+x^2} \cdot \frac{1}{\pi(1+x^2)} dx$$

$$V(Y) = E \left( \left( \frac{2x}{1+x^2} \right)^2 \right) = \int_{-\infty}^{\infty} \frac{4x^2}{(1+x^2)^3} dx$$

\[ \left[ \text{let, } x = \tan \theta \right. \]

\[ dx = \sec^2 \theta d\theta \]

\[ \int_{\frac{\pi}{2}}^{\frac{\pi}{2}} \]

\[ \frac{1}{\pi} \int_{0}^{\frac{\pi}{2}} \sec^2 \theta \cdot \sec^2 \theta d\theta \]

\[ \int_{0}^{\frac{\pi}{2}} \sin^2 \theta \cdot \cos^2 \theta d\theta \]

\[ \int_{0}^{\frac{\pi}{2}} \sin^2 \theta \cdot \cos^2 \theta d\theta \]

\[ \frac{1}{2} \int_{0}^{\frac{\pi}{2}} \sin^2 \theta \cdot \cos^2 \theta d\theta \]

\[ = \frac{1}{2} \left( \frac{3}{8} \right) \]

\[ = \frac{1}{2} \left( \frac{3}{8} \right) \]

\[ = \frac{1}{2} \left( \frac{3}{8} \right) \]

\[ = \frac{1}{2} \left( \frac{3}{8} \right) \]

\[ = 1 \]
\* Problem 7. Let \( X \sim C(\mu, \lambda) \) then find \( E\left( \frac{2}{3\mu} \log f(x) \right)^n \).

\[ f(x) = \frac{1}{\mu} \frac{1}{1 + (x - \mu)^2} \quad I_x (-\infty, \infty). \]

\[ \ln f(x) = -\ln \pi - \ln \left[ 1 + (x - \mu)^2 \right]. \]

\[ \frac{d}{dx} \ln f(x) = \frac{2(x - \mu)}{1 + (x - \mu)^2} = \gamma \quad \text{say}. \]

\[ E(Y) = \int_0^\infty \frac{2(x - \mu)^2}{1 + (x - \mu)^2} \cdot \frac{1}{\pi (1 + (x - \mu)^2)} \, dx \]

\[ = \frac{4}{\pi} \int_{-\infty}^{\infty} \frac{e^{-\gamma}}{(1 + \gamma)^3} \, d\gamma \quad [ \text{let}, \ x - \mu = 2 \]

\[ = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-\gamma}}{(1 + \gamma)^3} \, d\gamma \quad [ \text{see the previous problem}] \]

\[ = 1. \]

\* Problem 8. If \( X \sim C(\mu, \sigma) \) find the CDF and median of the distribution.

\[ F_X(x) = \int_0^x f_X(t) \, dt = \frac{1}{\pi} \int_0^x \frac{1}{\sqrt{t^2 + (x - \mu)^2}} \, dt \]

\[ = \frac{1}{\pi} \lim_{a \to \infty} \int_{-a}^{a} \frac{1}{1 + 2z} \, dz, \quad \text{where} \quad z = \frac{x - \mu}{\sigma} \]

\[ = \frac{1}{\pi} \lim_{a \to \infty} \int_{-a}^{a} \frac{1}{1 + 2z} \, dz \]

\[ = \frac{1}{\pi} \lim_{a \to \infty} \left[ \ln(1 + 2z) \right]_{-a}^{a} \]

\[ = \frac{1}{\pi} \lim_{a \to \infty} \left[ \ln(1 + 2a) - \ln(1 - 2a) \right] \]

\[ = \frac{1}{\pi} \left[ \ln(1 + \infty) - \ln(1 - \infty) \right] \]

\[ = \frac{1}{\pi} \left[ \ln(\infty) - \ln(0) \right] \]

\[ = \frac{1}{\pi} \left[ \infty - \ln(0) \right] \]

\[ = 1. \]

**Median:** The median of \( X \sim C(\mu, \sigma) \) is given by

\[ F_X(x) = \frac{1}{2} \]

\[ \Rightarrow \frac{1}{2} + \frac{1}{\pi} \text{tm}^{-1} \left( \frac{x - \mu}{\sigma} \right) = \frac{1}{2} \]

\[ \Rightarrow x = \mu \]

Hence, \( x = \mu \) is the median of \( C(\mu, \sigma) \) distribution.
Two dimensional random Variable

Two dimensional random vector: Consider a probability space \((\Omega, \mathcal{F}, P)\) arising out of a random experiment. A vector of functions \(\mathbf{X} = (X_1, X_2)\) which maps \(\Omega\) into \(\mathbb{R}^2\) is said to be two dimensional random vector, if for each \(\omega \in \Omega, \ i = 1, 2\,
\{ \omega : X_1(\omega) \leq x_1, X_2(\omega) \leq x_2 \} \in \mathcal{F}.

Example: An unbiased coin is tossed twice.

\[ \Omega = \{ HH, HT, TH, TT \} \]

We take \(\Gamma\) as a class of all subsets of \(\Omega\).

Let \(X_1\) denotes the number of heads obtained, and \(X_2\) denotes the number of tails obtained.

\[ \{ \omega : X_1(\omega) \leq x_1 \} = \begin{cases} \emptyset & \text{cohen } x_1 < 0 \\ \{ TT \} & \text{cohen } 0 \leq x_1 < 1 \\ \{ HH, HT, TT \} & \text{cohen } 1 \leq x_1 < 2 \\ \Omega & \text{cohen } x_1 \geq 2 \end{cases} \]

\[ \{ \omega : X_2(\omega) \leq x_2 \} = \begin{cases} \emptyset & \text{cohen } x_2 < 0 \\ \{ HH \} & \text{cohen } 0 \leq x_2 < 1 \\ \{ HH, TH, TT \} & \text{cohen } 1 \leq x_2 < 2 \\ \Omega & \text{cohen } x_2 \geq 2 \end{cases} \]

Now, \(\{ \omega : X_1(\omega) \leq x_1, X_2(\omega) \leq x_2 \} \in \mathcal{F}\) if \(X_1, X_2 \in \mathcal{F}\).

\[ \begin{align*}
\{ \omega : X_1(\omega) \leq x_1, X_2(\omega) \leq x_2 \} &= \emptyset, \text{ cohen } x_1 < 0, x_2 < 0 \\
&= \{ TT \}, \text{ cohen } 0 \leq x_1 < 1, 0 \leq x_2 < 1 \\
&= \emptyset, \text{ cohen } 0 \leq x_1 < 1, 1 \leq x_2 < 2 \\
&= \{ HH \}, \text{ cohen } 1 \leq x_1 < 2, 0 \leq x_2 < 1 \\
&= \emptyset, \text{ cohen } 1 \leq x_1 < 2, 1 \leq x_2 < 2 \\
&= \{ HH, HT, TT \}, \text{ cohen } 1 \leq x_1 < 2, x_2 \geq 2 \\
&= \emptyset, \text{ cohen } 1 \leq x_1 < 2, x_2 \geq 2 \\
&= \{ HH, TH, TT \}, \text{ cohen } x_1 \geq 2, 0 \leq x_2 < 1 \\
&= \emptyset, \text{ cohen } x_1 \geq 2, 1 \leq x_2 < 2 \\
&= \{ HH, TH, TT \}, \text{ cohen } x_1 \geq 2, x_2 \geq 2 \\
&= \Omega, \text{ cohen } x_1 \geq 2, x_2 \geq 2
\end{align*} \]

\(\mathbf{X} = (X_1, X_2)\) is a two dimensional random vector.
1. **Distribution function of \( x \)**

Let \( x = (x_1, x_2) \) be a two-dimensional random vector. The distribution function of \((x_1, x_2)\) is a function of \( F_{x_1, x_2}(x_1, x_2) \) such that:

\[
F_{x_1, x_2}(x_1, x_2) = \mathbb{P}(x_1 \leq x_1, x_2 \leq x_2) \quad \forall (x_1, x_2) \in \mathbb{R}^2.
\]

In the above example,

\[
F_{x_1, x_2}(x_1, x_2) = \begin{cases} 
0 & \text{cohen } x_1 < 0, x_2 \in \mathbb{R} \\
0 & \text{cohen } x_2 < 0, x_1 \in \mathbb{R} \\
0 & \text{cohen } 0 \leq x_1 < 1, 0 \leq x_2 < 1 \\
0 & \text{cohen } 0 \leq x_1 < 1, 1 \leq x_2 \leq 2 \\
1/4 & \text{cohen } 1 \leq x_1 < 2, 0 \leq x_2 \leq 1 \\
3/4 & \text{cohen } 1 \leq x_1 < 2, 1 \leq x_2 < 2 \\
1 & \text{cohen } x_1 \geq 2, x_2 \geq 2 \end{cases}
\]

\[
F_{x_1, x_2}(x_1, \infty) = \mathbb{P}(x_1 \leq x_1) = F_{x_1}(x_1) = \text{Marginal Distribution Function of } x_1.
\]

\[
F_{x_1, x_2}(\infty, x_2) = \mathbb{P}(x_2 \leq x_2) = F_{x_2}(x_2) = \text{Marginal Distribution Function of } x_2.
\]

\[
F_{x_1, x_2}(\infty, \infty) = 1
\]

\[
F_{x_1, x_2}(-\infty, x_2) = F_{x_1}(x_1) = F_{x_1}(x_1, -\infty)
\]

**Problem 1:**

If \( A = F_{x_1}(x_1) + F_{x_2}(x_2) \), \( G_1 = \sqrt{F_{x_1}(x_1) F_{x_2}(x_2)} \),

Then show that, \( 2A - 1 \leq F_{x_1, x_2}(x_1, x_2) \leq G_1 \).

**Solution:**

\[
2A - 1 = F_{x_1}(x_1) + F_{x_2}(x_2) - 1
\]

\[
= F_{x_1}(x_1) - \mathbb{P}(x_2 > x_2)
\]

\[
= \mathbb{P}(x_1 \leq x_1) - \mathbb{P}(x_2 > x_2)
\]

\[
F_{x_1, x_2}(x_1, x_2) = \mathbb{P}(x_1 \leq x_1, x_2 \leq x_2), \text{ let } C = x_1 \leq x_1 \& D = x_2 \leq x_2,
\]

\[
p(C) + p(D) - 1 = p(C \cap D)
\]

\[
+ 2A - 1 = F_{x_1, x_2}(x_1, x_2)
\]

\[
p(C \cap D) \leq p(C) \leq \{p(C \cap D)\} \leq p(C) = p(D)
\]

\[
F_{x_1, x_2}(x_1, x_2) \leq \sqrt{F_{x_1}(x_1) F_{x_2}(x_2)} = G_1
\]

\[
F_{x_1, x_2}(x_1, x_2) \leq G_1
\]

\[
2A - 1 = F_{x_1, x_2}(x_1, x_2) \leq G_1
\]
Result: Necessary & sufficient conditions for a function to be joint distribution function.

A Function \( F_{X_1, X_2}(x_1, x_2) \) is the joint distribution function of some 2-dimensional random variable \( (X_1, X_2) \) \iff (if and only if)

\[
\Delta F(x_1, x_2) = F(x_1 + h_1, x_2 + h_2) - F(x_1, x_2) - F(x_1 + h_1, x_2) + F(x_1, x_2 + h_2) \\
\geq 0 ;
\]

\[
F(-\infty, x_2) = F(x_1, -\infty) = 0 ;
\]

\[
F(x_1, -\infty) = 1 ;
\]

\[
F(x_1 + 0, x_2) = F(x_1, x_2 + 0) = F(x_1, x_2) .
\]

Proof: Note that,

\[
\Delta F(x_1, x_2) = F(x_1 + h_1, x_2 + h_2) - F(x_1, x_2) - F(x_1 + h_1, x_2) + F(x_1, x_2 + h_2) \\
= \sum \frac{\partial^2 F}{\partial x_1 \partial x_2} + \frac{\partial F}{\partial x_1} \frac{\partial F}{\partial x_2}.
\]

Also, note that the probability for the rectangle

\[
\mathbb{P}(\{x_1 \leq x \leq x_1 + h_1, x_2 \leq x_2 \leq x_2 + h_2 \})
\]

which necessarily belongs to \( \mathbb{B} \), equals the expression on the right hand side. Hence,

\[
\Delta F(x_1, x_2) = \sum \frac{\partial^2 F}{\partial x_1 \partial x_2} + \frac{\partial F}{\partial x_1} \frac{\partial F}{\partial x_2} > 0 .
\]

Denote by \( A_{n_1, n_2} \) the measurable set

\[
[-\infty, x \leq n_1, -\infty, x \leq n_2] ,
\]

where \( n \) is a positive integer. For fixed \( x_2 \), the sequence \( \{A_{n_1, x_2}\} \) is a contracting sequence whose limit is \( \emptyset \). It follows that

\[
\lim F(-\infty, x_2) = \lim P(A_{n_1, x_2}) = P(\lim A_{n_1, x_2}) = P(\emptyset) = 0 ,
\]

i.e., \( F(-\infty, x_2) = 0 \).

In a similar way, we have

\[
F(x_1, -\infty) = 0 .
\]
Consider the sets,
\[ A_n = \left[ -\alpha < x_1 \leq \alpha + \frac{1}{n}, \ -\alpha < x_2 \leq \alpha \right], \]
for positive integers \( n \). Now, if \( A_n \) is an expanding sequence of measurable sets whose limit is \( \mathbb{R} \), hence 
\[ \lim F(m, n) = \lim P(A_n) = P(\lim A_n) = P(\mathbb{R}) = 1, \]
i.e., \( F(\infty, \infty) = 1 \).

Let \( C_n = \left[ -\alpha < x_1 \leq \alpha + \frac{1}{n}, \ -\alpha < x_2 \leq \alpha \right] \),
where \( n \) is a positive integer, for fixed \( \alpha_1, \alpha_2 \); \( C_n \) is a contracting sequence of measurable sets and 
\[ \lim C_n = \left[ -\alpha < x_1 \leq \alpha_1, \ -\alpha < x_2 \leq \alpha_2 \right]. \]
Hence,
\[ \lim P(C_n) = P(\lim C_n) = P\left[ -\alpha < x_1 \leq \alpha_1, \ -\alpha < x_2 \leq \alpha_2 \right], \]
i.e., \[ \lim_{n \to \infty} F(x_1 + \frac{1}{n}, x_2) = F(x_1, x_2). \]
But,
\[ \lim_{\epsilon \to 0} F(x_1 + \epsilon, x_2) = \lim F\left( x_1 + \frac{1}{n}, x_2 \right) \]
and since \( F(x_1 + 0, x_2) \) is, by definition, the same as
\[ \lim_{\epsilon \to 0} F(x_1 + \epsilon, x_2), \] we have:
\[ F(x_1 + 0, x_2) = F(x_1, x_2). \]

By a similar argument, we get
\[ F(x_1, x_2 + 0) = F(x_1, x_2). \]

Note: \[ P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] \]
\[ = F(\alpha_1 + \alpha, \alpha_2 + \alpha) - F(\alpha_1 + \alpha, \alpha_2) - F(\alpha_1, \alpha_2 + \alpha) + F(\alpha_1, \alpha_2) \]
\[ \text{Proof:} \]
\[ P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] \]
\[ = P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] - P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] \]
\[ = P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] - P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] + P\left[ \alpha_1 < x_1 \leq \alpha_1 + \alpha, \ -\alpha_2 < x_2 \leq \alpha_2 + \alpha \right] \]
\[ = F(\alpha_1 + \alpha, \alpha_2 + \alpha) - F(\alpha_1, \alpha_2 + \alpha) - F(\alpha_1 + \alpha, \alpha_2) + F(\alpha_1, \alpha_2). \]
Problem 2. Show that the function \( F(\alpha_1, \alpha_2) \) is not a distribution function, where:
\[
F(\alpha_1, \alpha_2) = \begin{cases} 
0 & \text{if } \alpha_1 + \alpha_2 \leq 1 \\
1 & \text{if } \alpha_1 + \alpha_2 > 1 
\end{cases}
\]

**Solution:** Let us take \((\alpha_1, \alpha_2) = (0, 0)\) and \((\alpha_1 + \alpha_2) = (1, 1)\).
\[
\therefore F(\alpha_1 + \alpha_1, \alpha_2 + \alpha_2) = F(\alpha_1 + \alpha_1, \alpha_2) - F(\alpha_1, \alpha_2 + \alpha_2) + F(\alpha_1, \alpha_2)
\]
\[
= F(1, 1) - F(0, 0) - F(0, 1) + F(0, 0)
\]
\[
= 1 - 1 - 1 + 1
\]
\[
= 0
\]

So, \( F(\alpha_1, \alpha_2) \) does not satisfy the property of non-negativity for bivariate distribution.

Hence, \( F(\alpha_1, \alpha_2) \) is not a distribution function.

**Result:** If \( F_1(\alpha_1) \) and \( F_2(\alpha_2) \) are univariate distribution functions, then the function \( F(\alpha_1, \alpha_2) \) defined by
\[
F(\alpha_1, \alpha_2) = F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) \left[ 1 + \Theta (1 - F_{X_1}(\alpha_1)) (1 - F_{X_2}(\alpha_2)) \right], \Theta \geq 1
\]
is a joint distribution function.

**Proof:** \( F(\alpha_1, \alpha_2) = F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) \left[ 1 + \Theta (1 - F_{X_1}(\alpha_1)) (1 - F_{X_2}(\alpha_2)) \right] \)
\[
= F(-\infty, \alpha_2) = 0 = F(\alpha_1, -\infty)
\]

Now, let,
\[
G_{1X_1}(\alpha_1) = (1 - F_{X_1}(\alpha_1)) F_{X_1}(\alpha_1)
\]
\[
G_{1X_2}(\alpha_2) = (1 - F_{X_2}(\alpha_2)) F_{X_2}(\alpha_2)
\]
\[
\therefore F(\alpha_1, \alpha_2) = F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) + \Theta G_{1X_1}(\alpha_1) G_{1X_2}(\alpha_2)
\]
\[
\Delta_2 F(\alpha_1, \alpha_2) = F(\alpha_1 + \alpha_1, \alpha_2 + \alpha_2) - F(\alpha_1 + \alpha_1, \alpha_2) - F(\alpha_1, \alpha_2 + \alpha_2) + F(\alpha_1, \alpha_2)
\]
\[
\Delta_1 F_{X_1}(\alpha_1) = F_{X_1}(\alpha_1 + \alpha_1) - F_{X_1}(\alpha_1) + G_{1X_1}(\alpha_1)
\]
\[
\Delta_2 F_{X_2}(\alpha_2) = F_{X_2}(\alpha_2 + \alpha_2) - F_{X_2}(\alpha_2) + G_{1X_2}(\alpha_2)
\]
\[
\Delta_2 F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) + \Theta \Delta_2 G_{1X_1}(\alpha_1) G_{1X_2}(\alpha_2)
\]

**Note:** \( \Delta_2 F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) \)
\[
= F_{X_1}(\alpha_1 + \alpha_1) F_{X_2}(\alpha_2 + \alpha_2) + F_{X_1}(\alpha_1) F_{X_2}(\alpha_2) - F_{X_1}(\alpha_1 + \alpha_1) F_{X_2}(\alpha_2)
\]
\[
= F_{X_1}(\alpha_1) F_{X_2}(\alpha_2 + \alpha_2) - F_{X_2}(\alpha_2)
\]
\[
= \Delta_1 F_{X_1}(\alpha_1) \Delta_2 F_{X_2}(\alpha_2)
\]
\[ A_2 G_{X_1}(x_1) G_{X_2}(x_2) = A_1 G_{X_1}(x_1) G_{X_2}(x_2) \]
\[ = A_1 F_{X_1}(x_1)(1 - F_{X_2}(x_2)) A_1 F_{X_2}(x_2)(1 - F_{X_2}(x_2)) \]
\[ = A_1 F_{X_1}(x_1)(1 - F_{X_1}(x_1)) A_1 F_{X_2}(x_2)(1 - F_{X_2}(x_2)) \]

\[ A_1 G_{X_1}(x_1) = G_{X_1}(x_1 + h_1) - G_{X_1}(x_1) \]
\[ = F_{X_1}(x_1 + h_1)(1 - F_{X_1}(x_1)) - F_{X_1}(x_1)(1 - F_{X_1}(x_1)) - \left( F_{X_1}(x_1 + h_1) - F_{X_1}(x_1) \right) \]
\[ = A_1 F_{X_1}(x_1) \left\{ F_{X_1}(x_1 + h_1) + F_{X_1}(x_1) \right\} \]
\[ - A_1 F_{X_1}(x_1) \left\{ 1 - F_{X_1}(x_1 + h_1) - F_{X_2}(x_1) \right\} \]
\[ \left\{ F_{X_1}(x_1 + h_1) - F_{X_1}(x_1) \right\} \left\{ 1 - F_{X_2}(x_1 + h_1) - F_{X_2}(x_1) \right\} \]

\[ A_1 G_{X_2}(x_2) = \left\{ 1 - F_{X_2}(x_2 + h_2) - F_{X_2}(x_2) \right\} \left\{ 1 - F_{X_2}(x_2 + h_2) - F_{X_2}(x_2) \right\} \]

\[ \ast = A_1 F_{X_1}(x_1) A_1 F_{X_2}(x_2) + \Theta A_1 F_{X_1}(x_1) A_1 F_{X_2}(x_2) \]
\[ \left\{ 1 - F_{X_1}(x_1 + h_1) - F_{X_1}(x_1) \right\} \left\{ 1 - F_{X_2}(x_2 + h_2) - F_{X_2}(x_2) \right\} \]

Note, marginal distribution function of \( X_1, X_2 \) is

\[ P[X_1 \leq x_1] = F(x_1, x_2) = F_{X_1}(x_1) \]
\[ P[X_2 \leq x_2] = F(x_1, x_2) = F_{X_2}(x_2) \]

Given two variable marginal distribution function, one can generate a joint distribution function with some marginals. It is due to E. J. Gumbel.
**Independence**

- **Definition:** The random variables $X_1$ and $X_2$ are said to be independent if
  \[ F_{X_1X_2}(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2) \quad \forall (x_1, x_2) \in \mathbb{R}^2 \]

**Problem 3.** Find the marginal distribution function of $X_1, X_2$ and check whether they are independent or not.

Given:
\[ F_{X_1X_2}(x_1, x_2) = 1 - e^{-x_1} - e^{-x_2} - e^{-x_1-x_2} \quad ; \quad x_1, x_2 > 0 \]

**Solution:**
- The marginal distribution of $X_1$ is:
  \[ F_{X_1}(x_1) = F_{X_1X_2}(x_1, \infty) = 1 - e^{-x_1} \quad [\because e^{-\infty} = 0] \]
  The marginal distribution of $X_2$ is:
  \[ F_{X_2}(x_2) = 1 - e^{-x_2} = F_{X_1X_2}(\infty, x_2) \quad [\because x_1 \sim \text{Exp}(0, 1)] \]

Two RV's are said to be independently distributed if
\[ F_{X_1X_2}(x_1, x_2) = F_{X_1}(x_1)F_{X_2}(x_2) \]

Here,
\[ F_{X_1}(x_1) = 1 - e^{-x_1} \quad ; \quad F_{X_2}(x_2) = 1 - e^{-x_2} \]
\[ F_{X_1X_2}(x_1, x_2) = 1 - e^{-x_1} - e^{-x_2} + e^{-x_1-x_2} = F_{X_1}(x_1)F_{X_2}(x_2) \]

\[ \therefore \text{They are independent.} \]

**Problem 4.** \[ f(x, y) = 2e^{-x-y} \quad \text{for} \quad 0 < x, y < \infty \]

Find the marginal PDF of $X$ & $Y$.

**Solution:**
- The marginal PDF of $y$ is given by:
  \[ f_{Y}(y) = \int_{0}^{\infty} f(x, y) \, dx = \int_{0}^{\infty} 2e^{-x-y} \, dx = 2e^{-y} \left[ -e^{-x} \right]_{0}^{\infty} = 2e^{-y} \]

- The marginal PDF of $x$ is given by:
  \[ f_{X}(x) = \int_{0}^{\infty} f(x, y) \, dy = 2e^{-x} \int_{0}^{\infty} e^{-y} \, dy = 2e^{-x} \left[ -e^{-y} \right]_{0}^{\infty} = 2e^{-x} \]

- **Definition:** The random variables $X_1, X_2$ are said to be independent iff,
  \[ f_{X_1X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2) \quad \forall (x_1, x_2) \]
i) **Discrete Random Vector:** If a two-dimensional random variable \((x_1, x_2)\) takes only finite or countable infinite number of pairs of values \((x_1, x_2) \in \mathbb{R}^2\), then the random variable \(X\) is of discrete type. So, if a countable set \(c \subset \mathbb{R}\)

\[
P\left[(x_1, x_2) \in c^2\right] = 1.
\]

Consider a function \(f(x_1, x_2) \in \mathbb{R}\)

\[
f(x_1, x_2) = \begin{cases} 
P[X_1 = x_1, X_2 = x_2], & \text{cohere,} \\ 0, & \text{otherwise.} \end{cases}
\]

The function \(f(x_1, x_2)\) is called the PMF of \((x_1, x_2)\) if it satisfies the following conditions:

1. \(f(x_1, x_2) \geq 0 \forall (x_1, x_2) \in \mathbb{R}^2\)
2. \(\sum_{x_2} f(x_1, x_2) = 1\)

The marginal PMF of \(X_1\) is given by,

\[
f_{X_1}(x_1) = \sum_{x_2} f(x_1, x_2)
\]

\[
f_{X_2}(x_2) = \sum_{x_1} f(x_1, x_2)
\]

**Example:** Consider the bivariate table:

<table>
<thead>
<tr>
<th>(X)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.1</td>
<td>0.2</td>
<td>0.1</td>
<td>0.4</td>
</tr>
<tr>
<td>2</td>
<td>0.3</td>
<td>0.1</td>
<td>0.1</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.4</td>
<td>0.4</td>
<td>0.2</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**Theorem:** If \(X_1, X_2\) are jointly distributed in the discrete form with \(f(x_1, x_2)\) as the joint pmf, then the marginal p.m.f.s. 

**Marginal Probability of \(X_1\):**

\[
P(X = 1) = 0.4 \\
P(X = 2) = 0.4 \\
P(X = 3) = 0.2
\]

**Marginal Probability of \(Y_1\):**

\[
P(Y = 1) = 0.4 \\
P(Y = 2) = 0.5 \\
P(Y = 3) = 0.1
\]
Problem 5. Verify whether the following function is a joint PMF

\[ f(x, y) = \frac{e^{-2}}{x! (y-x)!}, \quad x=0, 1, 2, \ldots, y \]

Then find the marginal PMF of \( X \) & \( Y \). Check \( X \) & \( Y \) are independent or not.

**Soln.**

\[ f(x, y) > 0 \quad \forall \ (x, y) \]

\[ \begin{align*}
  \frac{2^x}{y^y} \cdot \frac{1}{x! (y-x)!} \\
  = e^{-2} \sum_{y=0}^{\infty} \frac{1}{y^y} \sum_{x=0}^{y} \frac{1}{x! (y-x)!} \\
  = e^{-2} \sum_{y=0}^{\infty} \frac{2^y}{y^y} \\
  = e^{-2} \cdot 2^1 = 1
\end{align*} \]

\[ \therefore f(x, y) \text{ is a joint PMF.} \]

\[ f_X(x) = \sum_{y=x}^{\infty} \frac{e^{-2}}{x! (y-x)!} \]

\[ = e^{-2} \sum_{y=x}^{\infty} \left( \frac{1}{y^y} \sum_{x=0}^{y} \frac{1}{x! (y-x)!} \right) \\
= e^{-2} \cdot \frac{1}{x!} \]

\[ = \frac{e^{-1}}{x!}, \quad x=0, 1, 2, \ldots, \infty \]

\[ X \sim \text{Poi}(1), \]

\[ f_Y(y) = \frac{1}{y!} \cdot \frac{e^{-2}}{(y-x)!} \]

\[ = \frac{e^{-2} \cdot 2^y}{y!}, \quad y=0, 1, \ldots, \infty \]

\[ \therefore Y \sim \text{Poi}(2). \]

Here, \( f_X(x) f_Y(y) \neq f_{XY}(x, y) \)

\[ \Rightarrow X, Y \text{ are not independent.} \]
Covariance: The covariance between $X_1$ and $X_2$ is defined by
\[ \text{cov}(X_1, X_2) = E(X_1X_2) - E(X_1)E(X_2) \]

The correlation coefficient between $X_1$ and $X_2$ is defined by
\[ p_{X_1X_2} = \frac{\text{cov}(X_1, X_2)}{\sqrt{\text{var}(X_1)\text{var}(X_2)}} \]

\[ -1 \leq p_{X_1X_2} \leq 1 \]

Problem 5. (Continuation of Problem 5)
Find the $E(X) = E(Y) = 2$
$I = E(X) = 1, \text{var}(Y) = 2$

\[ E(Y) = \sum_{y=0}^{\infty} \sum_{x=0}^{\infty} x y e^{-2} \frac{x! y!}{(x+y)!} \]
\[ = e^{-2} \sum_{y=0}^{\infty} \frac{y!}{y!} \sum_{x=0}^{\infty} \frac{x^2}{(x+1)!} e^{-x} \]
\[ = e^{-2} 2 \sum_{y=0}^{\infty} \frac{y!}{y!} \sum_{x=0}^{\infty} \frac{x^{y+1}}{(y+1)!} e^{-x} \]
\[ = e^{-2} \sum_{y=0}^{\infty} \frac{y!}{y!} \sum_{x=0}^{\infty} \frac{x^{y+1}}{(y+1)!} e^{-x} \]
\[ = \frac{1}{2} E(Y^2), \text{ hence } Y \sim \text{Poi}(2) \]
\[ = \frac{1}{2} \left[ 2^2 + 2 \right] \]
\[ = 3 \]

\[ \text{cov}(X,Y) = E(XY) - E(X)E(Y) \]
\[ = 3 - 1 \times 2 \]
\[ = 1 \]

\[ p_{XY} = \frac{1}{\sqrt{1 \times 2}} \]
\[ = \frac{1}{\sqrt{2}} \]
Continuous Random Vector: A two-dimensional random variable \( X_1, X_2 \) is said to be continuous if \( F_{X_1 X_2}(x_1, x_2) \) is everywhere continuous on \( \mathbb{R}^2 \).

Absolutely Continuous Random Vector: A two-dimensional continuous random variable \( X_1, X_2 \) is said to be absolutely continuous if \( \exists \) a non-negative integrable function \( f_{X_1 X_2}(x_1, x_2) \in \mathbb{R}^2 \):

\[
F_{X_1 X_2}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_1 X_2}(u, v) \, du \, dv \quad \forall \ x_1, x_2 \in \mathbb{R}.
\]

In that case, \( \frac{d}{dx_1} \frac{d}{dx_2} F_{X_1 X_2}(x_1, x_2) = f_{X_1 X_2}(x_1, x_2) \) exists if the derivative exists.

The function \( f_{X_1 X_2}(x_1, x_2) \) is called the joint PDF of \( X_1, X_2 \) if it satisfies the following two conditions:

\[ f_{X_1 X_2}(x_1, x_2) > 0 \quad \forall \ x_1, x_2 \in \mathbb{R} \]

\[ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) \, dx_1 \, dx_2 = 1. \]

Marginal PDF:

\[
f_{X_1}(x_1) = \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) \, dx_2 = g(x_1)
\]

\[
f_{X_2}(x_2) = \int_{-\infty}^{\infty} f_{X_1 X_2}(x_1, x_2) \, dx_1 = h(x_2)
\]

These two are called the marginal PDF of \( X_1, X_2 \).

Problem 7. Let \( g \) and \( h \) be 2 PDFs with corresponding distribution function \( G, H \). Consider the function

\[
f_{X Y}(x, y) = g(x) h(y) [1 + \alpha (2G(x) - 1) (2H(y) - 1)]
\]

Show that \( f \) is a joint PDF with the given marginal PDFs \( g, h \).

Solution:

\[ 0 < G(x) < 1 \]
\[ -1 < 2G(x) - 1 < 1 \]
\[ |2G(x) - 1| < 1 \]
\[ |x| \leq 1 \quad \text{and} \quad |2H(y) - 1| < 1 \]
\[ -1 \leq \alpha (2G(x) - 1) (2H(y) - 1) \leq 1 \]
\[ 0 \leq [1 + \alpha (2G(x) - 1) (2H(y) - 1)] \leq 2 \]
\[ \Rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X Y}(x, y) > 0 \]
Now, \[ \int \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx \, dy \]
\[ = \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} g(x) \, dx \right] \, dy \]
\[ = 1 + \alpha \int_{0}^{a} \frac{u}{2} \, du \times \int_{0}^{\frac{a}{2}} \frac{v}{2} \, dv \]
\[ = 1. \]

**Marginal PDFs**

\[ f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dy \]
\[ = \int_{0}^{\infty} g(x) \, dy + \alpha \int_{0}^{\infty} g(x) \, dy \]
\[ = g(x) \]

\[ f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx = h(y) \]

**Problem.** Let \( X \) and \( Y \) be such that \( Y = 2X \). Show that

\[ F_{X,Y}(x,y) = \begin{cases} 1, & \text{if } x \leq \alpha \leq \frac{y}{2} \\ 0, & \text{otherwise} \end{cases} \]

**Solution.**

\[ F_{X,Y}(x,y) = P[X \leq x, Y \leq y] \]
\[ = P[X \leq \alpha, 2x \leq y] \]
\[ = P[X \leq \alpha, X \leq \frac{y}{2}] \]

Hence,
\[ F_{X,Y}(x,y) = g_X(x) \text{ if } x \leq \frac{y}{2} \]
\[ = g_X(\frac{y}{2}) \text{ if } x > \frac{y}{2} \]
**Conditional Distribution:**

- **Discrete Case:** Let \((x_1, x_2)\) be a 2-dimensional discrete random variable. The joint PMF \(f_{x_1,x_2}(x_1, x_2)\), the conditional PMF of \(x_1\) given \(x_2 = x_2\) is given by:

\[
\frac{f_{x_1|x_2}(x_1 | x_2)}{f_{x_2}(x_2)} = \frac{f_{x_1,x_2}(x_1, x_2)}{f_{x_2}(x_2)} \quad \text{if} \quad f_{x_2}(x_2) > 0
\]

Similarly, the conditional PMF of \(x_2\) given \(x_1 = x_1\) is given by:

\[
f_{x_2|x_1}(x_2 | x_1) = \frac{f_{x_1,x_2}(x_1, x_2)}{f_{x_1}(x_1)} \quad \text{if} \quad f_{x_1}(x_1) > 0
\]

If \(x_1\) and \(x_2\) are independent,

\[
f_{x_1|x_2}(x_1 | x_2) = f_{x_1}(x_1) \quad \text{if} \quad f_{x_2}(x_2) > 0
\]

\[
f_{x_2|x_1}(x_2 | x_1) = f_{x_2}(x_2) \quad \text{if} \quad f_{x_1}(x_1) > 0
\]

\[
f_{x_1,x_2}(x_1, x_2) = f_{x_1}(x_1) f_{x_2}(x_2)
\]

\[\star \text{ Problem 9. (Continuation of Problem 5.)} \]

Find the conditional PMF \(f_{x|y}(x|y)\) and \(f_{y|x}(y|x)\).

**Solution:**

\[
f_{x|y}(x|y) = \frac{P[X = x, Y = y]}{P[Y = y]} = \frac{e^{-\alpha} \alpha^y}{y!}
\]

\[
f_{y|x}(y|x) = \frac{P[X = x, Y = y]}{P[X = x]} = \frac{e^{-\alpha} \alpha^x}{x!}
\]

\[
f_{y|X}(y|x) = \frac{e^{-\alpha} \alpha^x}{y!}
\]

\[
f_{X|Y}(X|y) = \frac{e^{-\alpha} \alpha^y}{y!}
\]
Problem 11. A 2-dimensional random vector \((X_1, X_2)\) has

$$\text{PMF, } f_{X_1 X_2}(x_1, x_2) = \frac{n! \ p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{(n-x_1-x_2)}}{x_1! x_2! (n-x_1-x_2)!}$$

where \(0 \leq x_1 + x_2 \leq n\), \(-\text{binomial distribution}\)

- \(0 < p_1, p_2 < 1\) \(\Rightarrow\) \(p_1 + p_2 < 1\).

Find the marginal distribution of \(X_1, X_2\) and conditional distr. of \(X_1 | X_2 = x_2\).

Solution: The marginal PMF of \(X_1\) is,

$$f_{X_1}(x_1) = \sum_{x_2=0}^{n-x_1} \frac{n! \ p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{(n-x_1-x_2)}}{x_1! x_2! (n-x_1-x_2)!}$$

$$= \binom{n}{x_1} p_1^{x_1} (1-p_1-p_2)^{(n-x_1)} \sum_{x_2=0}^{n-x_1} p_2^{x_2} \frac{1}{x_2!}$$

$$= \binom{n}{x_1} p_1^{x_1} (1-p_1-p_2)^{(n-x_1)} \binom{n-x_1}{x_2} (p_2 + ! -p_1-p_2)^{(x_2-x_1)}$$

$$= \binom{n}{x_1} p_1^{x_1} (1-p_1)^{(x_2-x_1)} I_{x_1} (0,1,\ldots,n)$$

\(\therefore\) \(X_1 \sim \text{Bin}(n, p_1)\)

Similarly,

$$f_{X_2}(x_2) = \binom{n}{x_2} p_2^{x_2} (1-p_2)^{(n-x_2)} \sum_{x_1=0}^{n-x_2} p_1^{x_1} \frac{1}{x_1!}$$

$$= \binom{n}{x_2} p_2^{x_2} (1-p_2)^{(n-x_2)} \binom{n-x_2}{x_1} (p_1 + ! -p_1-p_2)^{(x_1-x_2)}$$

$$= \binom{n}{x_2} p_2^{x_2} (1-p_2)^{(n-x_2)} \binom{n-x_2}{x_1} (1-p_1-p_2)^{(x_1-x_2)}$$

\(\therefore\) \(X_2 \sim \text{Bin}(n, p_2)\)

Conditional PMF of \(X_1 | X_2 = x_2\) is,

$$f_{X_1/X_2}(x_1 | x_2) = \frac{f_{X_1 X_2}(x_1, x_2)}{f_{X_2}(x_2)}$$

$$= \frac{n! \ p_1^{x_1} p_2^{x_2} (1-p_1-p_2)^{(n-x_1-x_2)}}{x_1! x_2! (n-x_1-x_2)!} \binom{n-x_2}{x_1} (1-p_1-p_2)^{(x_1-x_2)}$$

$$= \binom{n-x_2}{x_1} (p_1 + ! -p_1-p_2)^{(x_1-x_2)}$$

\(\therefore\) \(X_1 | X_2 = x_2 \sim \text{Bin}(n-x_2, p_1)\), and \(0 < x_1 < n-x_2\) \(\text{ and } 0 < p_1 < 1\).
Absolutely Continuous Case: Let \( f_{X_1 X_2} \) be the joint PDF of \( X_1 \) and \( X_2 \) and let \( g_{X_1} \) be the marginal PDF of \( X_1 \), then for \( \alpha, \beta > 0 \),

\[
P\left[ \frac{\alpha - \beta}{2} \leq X_2 \leq \frac{\alpha + \beta}{2} \mid \frac{\alpha - \beta}{2} \leq X_1 \leq \frac{\alpha + \beta}{2} \right]
\]

\[
= \frac{\int_{\frac{\alpha - \beta}{2}}^{\frac{\alpha + \beta}{2}} \int_{\frac{\alpha - \beta}{2}}^{\frac{\alpha + \beta}{2}} f_{X_1 X_2}(u, v) \, du \, dv}{\int_{\frac{\alpha - \beta}{2}}^{\frac{\alpha + \beta}{2}} \int_{\frac{\alpha - \beta}{2}}^{\frac{\alpha + \beta}{2}} g_{X_1}(u) \, du \cdot \frac{g_{X_1}(x)}{g_{X_1}(x)}}
\]

\[
= \int_{\frac{\alpha - \beta}{2}}^{\frac{\alpha + \beta}{2}} \frac{f_{X_1 X_2}(x, x)}{g_{X_1}(x)} \, dx
\]

Now, as \( \alpha, \beta \to 0 \),

\[
\lim_{\alpha, \beta \to 0} P\left[ \frac{\alpha - \beta}{2} \leq X_2 \leq \frac{\alpha + \beta}{2} \mid \frac{\alpha - \beta}{2} \leq X_1 \leq \frac{\alpha + \beta}{2} \right] = \frac{f_{X_1 X_2}(x_1, x_2)}{g_{X_1}(x_1)}
\]

The L.H.S. is denoted by \( f_{X_2 | X_1}(x_2 \mid x_1) \). This behaves like a univariate probability density function \( f_{X_2 | X_1}(x_2 \mid x_1) \geq 0 \) and

\[
\int_{-\infty}^{\infty} \frac{f_{X_1 X_2}(x_1, x_2)}{g_{X_1}(x_1)} \, dx_2 = 1
\]

The function \( f_{X_2 \mid X_1}(x_2 \mid x_1) \) given by \( f_{X_2 \mid X_1}(x_2 \mid x_1) = \frac{f_{X_1 X_2}(x_1, x_2)}{g_{X_1}(x_1)} \) will therefore be called the conditional PMF \( (X_2 \mid X_1 = x_1) \).
Statement: If \( X_1 \) and \( X_2 \) are independently distributed random variables such that \( E(X_1) \) and \( E(X_2) \) exist, then \( E(X_1 + X_2) \) exists and \( E(X_1 + X_2) = E(X_1) + E(X_2) \).

Proof:

Let \( X_1 \) and \( X_2 \) be independent random variables. Then,

\[
E(X_1 + X_2) = \sum_{x \in S_1} \sum_{y \in S_2} (x + y) f_{X_1}(x) f_{X_2}(y)
\]

where \( S_1 = \text{domain of } X_1 \) and \( S_2 = \text{domain of } X_2 \).

By independence,

\[
f_{X_1}(x) f_{X_2}(y) = f_{X_1}(x) f_{X_2}(y)
\]

Thus,

\[
E(X_1 + X_2) = \sum_{x \in S_1} \sum_{y \in S_2} (x + y) f_{X_1}(x) f_{X_2}(y)
\]

\[
= \sum_{x \in S_1} \sum_{y \in S_2} x f_{X_1}(x) f_{X_2}(y) + \sum_{x \in S_1} \sum_{y \in S_2} y f_{X_1}(x) f_{X_2}(y)
\]

\[
= \sum_{x \in S_1} \sum_{y \in S_2} x f_{X_1}(x) f_{X_2}(y) + \sum_{y \in S_2} \sum_{x \in S_1} y f_{X_1}(x) f_{X_2}(y)
\]

\[
= \sum_{x \in S_1} \sum_{y \in S_2} x f_{X_1}(x) f_{X_2}(y) + \sum_{y \in S_2} \sum_{x \in S_1} y f_{X_1}(x) f_{X_2}(y)
\]

\[
= \sum_{x \in S_1} x f_{X_1}(x) \sum_{y \in S_2} f_{X_2}(y) + \sum_{y \in S_2} y f_{X_2}(y) \sum_{x \in S_1} f_{X_1}(x)
\]

\[
= \sum_{x \in S_1} x f_{X_1}(x) \cdot E(X_2) + \sum_{y \in S_2} y f_{X_2}(y) \cdot E(X_1)
\]

\[
= E(X_1) \cdot E(X_2) + E(X_2) \cdot E(X_1)
\]

\[
= 2E(X_1)E(X_2)
\]

Hence, \( E(X_1 + X_2) = E(X_1) + E(X_2) \).
**Remark:** If $x_1$ and $x_2$ are independent and $G_1(x_1)$ and $G_2(x_2)$ are functions of $x_1$ and $x_2$ so $E(G_1(x_1))$ and $E(G_2(x_2))$ exist and expectation of their product exists and 

$$E[G_1(x_1)G_2(x_2)] = E[G_1(x_1)]E[G_2(x_2)]$$

**Problem:** If $x_1$ and $x_2$ are independent random variables so

$$E(x_i) = \mu_i, \quad i = 1, 2$$

$$\sigma_i = \sigma_i, \quad i = 1, 2$$

Then

$$E(x_1) = E(x_2) = \sigma_1 \sigma_2 + \sigma_1 \mu_2 + \sigma_2 \mu_1$$

Find $\rho(x_1x_2)$ and $\rho(x_1, x_2) = \text{correlation coefficient between } x_1 \text{ and } x_2$.

**Solution:**

$$\rho(x_1x_2) = \frac{E(x_1x_2) - E(x_1)E(x_2)}{\sqrt{E(x_1^2) - (E(x_1))^2} \sqrt{E(x_2^2) - (E(x_2))^2}}$$

$\rho(x_1x_2) = 0$ implies independent.

**Note:** If $x_1, x_2$ are independently distributed then they must be uncorrelated but the converse may not be true.

**Solution:**

If $x_1, x_2$ are independent then

$$E(x_1x_2) = E(x_1)E(x_2)$$

$$\text{Cov}(x_1, x_2) = 0$$

$$\mathbb{P}[x_1, x_2]; \quad \mathbb{P}[x_1] \mathbb{P}[x_2]$$
Only if Part: — Consider the following counter example:

Let \( X_1 \) be a discrete random variable which takes 3 values \(-1, 0, 1\) each with probability \( \frac{1}{3} \). Define \( X_2 = X_1^2 \). Then \( X_2 \) takes two values 0, 1 with probability \( \frac{1}{3}, \frac{2}{3} \), i.e., the joint distribution of \( X_1, X_2 \) is given in the following table:

<table>
<thead>
<tr>
<th>( X_1 )</th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1</td>
<td>0</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
</tr>
<tr>
<td>0</td>
<td>( \frac{1}{3} )</td>
<td>0</td>
<td>( \frac{1}{3} )</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{1}{3} )</td>
</tr>
<tr>
<td>Total</td>
<td>( \frac{1}{3} )</td>
<td>( \frac{2}{3} )</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
E(X_1) = 0 \\
E(X_2) = 0 \times \frac{1}{3} + 1 \times \frac{2}{3} = \frac{2}{3} \\
E(X_1X_2) = -\frac{1}{3} + \frac{1}{3} = 0 \\
E(X_1^2) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3} \\
E(X_2) = \frac{2}{3} \\
Y(X_1) = \frac{2}{3} \\
Y(X_2) = \frac{2}{3} - \frac{2}{3} = \frac{2}{3} \\
\text{Un correlated but not independent.}
\]

Consider the joint distribution of 2 random variables \( X \) & \( Y \) as follows:

<table>
<thead>
<tr>
<th>( X )</th>
<th>( Y_1 )</th>
<th>( Y_2 )</th>
<th>( P[X = x_i] )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>( P_{11} )</td>
<td>( P_{12} )</td>
<td>( P_{16} )</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>( P_{21} )</td>
<td>( P_{22} )</td>
<td>( P_{20} )</td>
</tr>
<tr>
<td>( \beta )</td>
<td>( P_{01} )</td>
<td>( P_{02} )</td>
<td>1</td>
</tr>
</tbody>
</table>

\[
P[X = \alpha_1, Y = y_1] = P_{11} \\
P[X = \alpha_1, Y = y_2] = P_{12} \\
P[X = \alpha_2, Y = y_1] = P_{21} \\
P[X = \alpha_2, Y = y_2] = P_{22} \\
P[X = \alpha_1] = P_{10}
\]
\[ E(X) = x_1 p_{10} + x_2 p_{20} \]
\[ E(XY) = x_1 y_1 p_{11} + x_1 y_2 p_{12} + x_2 y_1 p_{21} + x_2 y_2 p_{22} \]
\[ E(Y) = y_1 p_{01} + y_2 p_{02} \]

Let \(X, Y\) be uncorrelated,
\[ E(XY) = E(X)E(Y) \]
\[ \Rightarrow \alpha_1 y_1 p_{11} + \alpha_1 y_2 p_{12} + \alpha_2 y_1 p_{21} - \left[ (\alpha_1 p_{10} + \alpha_2 p_{20}) (y_1 p_{01} + y_2 p_{02}) \right] \]
\[ \Rightarrow \alpha_1 y_1 (p_{11} - p_{10} p_{01}) + \alpha_2 y_2 (p_{22} - p_{20} p_{02}) + \alpha_2 y_1 (p_{21} - p_{20} p_{01}) \]
\[ + \alpha_1 y_2 (p_{12} - p_{10} p_{02}) = 0 \quad \cdots \quad (1) \]

From the joint distribution table, we get,

\[ \begin{align*}
\text{i)} & \quad p_{12} = p_{10} - p_{11}
\quad p_{22} = 1 - p_{10}

\therefore & \quad p_{12} - p_{10} p_{02} = p_{10} - p_{11} - p_{10} (1 - p_{01})
\quad = - \left[ p_{11} - p_{10} p_{01} \right] \quad \cdots \quad (2)
\end{align*} \]

\[ \begin{align*}
\text{ii)} & \quad p_{21} = p_{01} - p_{11}
\quad p_{20} = 1 - p_{10}

\therefore & \quad p_{21} - p_{20} p_{01} = p_{01} - p_{11} - p_{01} (1 - p_{10})
\quad = - \left[ p_{11} - p_{10} p_{01} \right] \quad \cdots \quad (3)
\end{align*} \]

\[ \begin{align*}
\text{iii)} & \quad p_{22} - p_{02} p_{20}
\quad = p_{11} - p_{10} p_{01} \quad \cdots \quad (4)
\end{align*} \]

\( (1) \) reduces to

\[ \left( p_{11} - p_{10} p_{01} \right) \left[ \alpha_1 y_1 + \alpha_2 y_2 - \alpha_2 y_1 - \alpha_1 y_2 \right] = 0 \]
\[ \Rightarrow \left( p_{11} - p_{10} p_{01} \right) \left( \alpha_1 - \alpha_2 \right) \left( y_1 - y_2 \right) = 0 \]
\[ \Rightarrow \quad p_{11} = p_{10} p_{01} \]

From (2), (3), (4); we get,

\[ \begin{align*}
\text{P}_{12} & = P_{10} P_{02} \\
\text{P}_{21} & = P_{20} P_{01} \\
\text{P}_{22} & = P_{20} P_{02} \\
\therefore & \text{P}_{ij} = P_{i0} P_{0j} \quad \forall \quad i,j \\
\therefore & \text{X and Y are independent. [Proved]} 
\end{align*} \]
Problem 12. Let $X$ and $Y$ be independently distributed random variables. Find:

$$P[X = i] = P[Y = i] = \frac{1}{n} \quad \forall i = 1 \ldots n.$$ 

**Solution:**

$$P[X = Y] = \sum_{i=1}^{n} P[X = i] P[Y = i] = \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n}{n} = \frac{1}{n}.$$ 

$$P[X < Y] = \sum_{i=1}^{n} P[X < Y, Y = i] = \sum_{i=1}^{n} P[X < Y \mid Y = i] P[Y = i] = \sum_{i=1}^{n} \frac{i-1}{n} \cdot \frac{1}{n} = \frac{1}{n} \left[ \sum_{i=1}^{n} \frac{i}{n} - \sum_{i=1}^{n} \frac{1}{n} \right] = \frac{1}{n} \left[ \frac{n(n+1)}{2n} - \frac{n}{n} \right] = \frac{1}{n} \left[ \frac{n^2 + n - 2n}{2n} \right] = \frac{1}{n} \left[ \frac{n+1}{2} - \frac{1}{n} \right].$$

$$P[X > Y] = 1 - P[X = Y] - P[X < Y] = 1 - \frac{1}{n} - \frac{m-1}{2n} = \frac{2n - 2 - n + 1}{2n} = \frac{n - 1}{2n}.$$
Moments of the Conditional Distribution:

The conditional mean of \( x_1 / x_2 = x_2 \) is defined by:
\[
E \left[ x_1 / x_2 = x_2 \right] = \sum_{x_1} x_1 f_{x_1 / x_2} (x_1 / x_2) \quad \text{[Provided it exists]}
\]
\[
= \eta_{x_2}, \text{ say}
\]

The conditional mean of \( x_2 / x_1 = x_1 \) is defined by:
\[
E \left[ x_2 / x_1 = x_1 \right] = \sum_{x_2} x_2 f_{x_2 / x_1} (x_2 / x_1) \quad \text{[Provided it exists]}
\]
\[
= \eta_{x_1}, \text{ say}
\]

Conditional variance of \( x_1 / x_2 = x_2 \) is:
\[
V \left( x_1 / x_2 = x_2 \right) = E \left[ (x_1 - \eta_{x_2})^2 / x_2 = x_2 \right] - \eta_{x_2}^2 \quad \text{[Provided it exists]}
\]

Similarly,
\[
V \left( x_2 / x_1 = x_1 \right) = E \left[ (x_2 - \eta_{x_1})^2 / x_1 = x_1 \right] - \eta_{x_1}^2 \quad \text{[Provided it exists]}
\]

Problem 13: A random vector \((X, Y)\) has PMF
\[
f_{XY}(x, y) = \frac{(\alpha + \beta + k - 1)!}{\alpha! \beta! (k - 1)!} \frac{\alpha x \beta y}{(1 - \alpha - \beta - k)^k}, \quad 0 < \alpha, \beta < \infty, \quad k > 1, \text{ is an integer.}
\]

where \( 0 < \alpha, \beta < 1 \) and \( \alpha + \beta < 1 \).

Find marginal and conditional distribution.

Solution:
\[
f_X(x) = \sum_{y=0}^{\infty} f_{XY}(x, y)
\]
\[
= \sum_{y=0}^{(x+k-1)} \frac{(\alpha + \beta + k - 1)!}{\alpha! \beta! (k - 1)!} \frac{\alpha x \beta y}{(1 - \alpha - \beta - k)^k} \frac{\alpha (x+y+k-1)!}{\alpha y! (k-1)!}
\]
\[
= \frac{\alpha x}{(1 - \alpha - \beta - k)^k} \frac{\alpha x \beta y}{(1 - \alpha - \beta - k)^k} \frac{\alpha (x+y+k-1)!}{\alpha y! (k-1)!}
\]
\[
= \binom{x+k-1}{x} \frac{\alpha x \beta y}{(1 - \alpha - \beta - k)^k}, \quad 0 \leq x < \infty, \quad 0 < \alpha, \beta < 1, \quad k > 1, \text{ is an integer.}
\]

Similarly,
\[
f_Y(y) = \binom{y+k-1}{y} \frac{\beta y}{1 - \alpha} \frac{\beta y}{1 - \alpha} \frac{(1 - \alpha - \beta - k)^k}{(1 - \alpha - \beta - k)^k}, \quad 0 \leq y < \infty, \quad 0 < \alpha, \beta < 1, \quad k > 1, \text{ is an integer.}
\[ \frac{f_{y/x}(y/x)}{f_{x/y}(x/y)} = \frac{(x+y+k-1)! \frac{1}{x!y!(k-1)!} \frac{p_1^x p_2^y (1-p_1-p_2)^k}{x! (k-1)! \left( \frac{p_1}{1-p_1-p_2} \right)^k (1-p_1-p_2)^k}}{(x+k-1)! \frac{1}{y!(x+k-1)!} \frac{p_2^y (1-p_2)^{x+k}}{y!}} \]

Similarly,
\[ f_{x/y}(x/y) = \frac{(x+y+k-1)!}{y! (x+k-1)!} \frac{p_1^x (1-p_1)^{y+k}}{x!}, \quad 0 \leq y < x, \quad 0 < p_1 < 1 \]

\[ E(X_1) = E \left[ E \left( \frac{X_1}{X_2} \right) \right] = E \left[ \eta_{X_2} \right] \]

**Result:** If \( E(X_1/X_2 = x_2) \) exists for almost all values of \( x_2 \) (i.e., for all values of \( x_2 \)) \( \exists \int x_2 E(\eta_{X_2}) > 0 \) then,

**Solution:**
\[ E(X_1) = \sum_{x_1} \frac{x_1}{x_1} f_{X_1}(x_1) \]

\[ = \sum_{x_1} \frac{x_1}{x_1} \left( \sum_{x_2} f_{X_1X_2}(x_1, x_2) \right) \]

\[ = \sum_{x_1} \sum_{x_2} \frac{x_1}{x_2} f_{X_1X_2}(x_1, x_2) \]

\[ = \sum_{x_1} \sum_{x_2} \frac{x_1}{x_2} f_{X_2}(x_2) f_{X_1/X_2}(x_1/x_2) \]

\[ = \sum_{x_1} \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) \right] f_{X_2}(x_2) \]

\[ = E \left[ \eta_{X_2} \right] \]

\[ = E \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]

\[ = \sum_{x_2} \left[ \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]

\[ = E \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]

\[ = E \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]

\[ = E \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]

\[ = E \left[ \sum_{x_2} \frac{x_1}{x_2} f_{X_1/X_2}(x_1/x_2) f_{X_2}(x_2) \right] \]
Result: If \( E[X_1 / X_2 = x_2] \) and \( Y[X_1 / X_2 = x_2] \) exist for almost all values of \( x_2 \) then,
\[
V(X_1) = E[Y(X_1 / X_2)] + Y[E(X_1 / X_2)]
\]

**Soln.**

R.H.S.
\[
E[Y(X_1 / X_2)] = E[E(Y(X_1 / X_2) - \eta X_2)]
\]
\[
= E[E(X_1 Y / X_2)] - E(\eta X_2)
\]
\[
= E(X_1 Y) - E(\eta X_2)
\]
\[
Y[E(X_1 / X_2)] = Y(\eta X_2)
\]
\[
= E(\eta X_2) - E[Y(\eta X_2)]
\]
\[
= E(\eta X_2) - E(Y(X_1))
\]

Thus,
\[
E[Y(X_1 / X_2)] + Y[E(X_1 / X_2)]
\]
\[
= E(X_1 Y) - E(Y(X_1))
\]
\[
= Y(X_1) = L.H.S.
\]

**Note:** If \( E(X_2 / X_1 = x_1) \) and \( V(X_2 / X_1 = x_1) \) exist for all values of \( x_1 \),
\[
V(X_2) = E(Y(X_2 / X_1)) + Y(E(X_2 / X_1))
\]

**Problem 14.** \( X \sim P(\lambda) \) and \( Y / X = \xi \sim Poi(\theta X) \), s.t.
\[
\rho_{XY} = \frac{\theta X}{1 + \theta X}
\]

**Soln.**

\[
E(Y) = E[E(Y | X)]
\]
\[
= E(\lambda X) = \lambda \mu
\]
\[
Y(Y) = E(X Y) + Y(\lambda X)
\]
\[
= E(X Y) = \lambda \mu + \lambda \mu
\]
\[
E(\lambda X) = \lambda E(X) = \lambda \mu
\]
\[
E(X Y) = E[X E(Y / X)]
\]
\[
= E[X, \lambda X]
\]
\[
= \lambda [\mu + \lambda]
\]
\[
\rho_{XY} = \sqrt{\frac{\lambda [\mu + \lambda] - \lambda \mu}{\lambda (\lambda \mu + \lambda \mu)}}
\]
\[
= \sqrt{\frac{\lambda}{1 + \theta X}}.
\]
Problem 15. Let \( (X, Y) \) be a bivariate discrete random variable.

\[
\begin{aligned}
    f_{X,Y}(x, y) &= \begin{cases} 
        \frac{1}{\binom{m+1}{2}} & \text{if } y = 1, 2, \ldots, x \\
        0 & \text{otherwise}
    \end{cases} \\
    m &\text{ a positive integer } > 1.
\end{aligned}
\]

Find \( E(X) \) and \( E(X/Y) \) and \( E[E(X/Y)] \).

**Solution:**

\[
\begin{aligned}
    f_X(x) &= \sum_{y=1}^{x} f_{X,Y}(x, y) = \frac{x}{\binom{m+1}{2}} = \frac{x}{\binom{m+1}{2}} \cdot \frac{m(m+1)(2m+1)}{2m+1} \\
    E(X) &= \sum_{x=1}^{m} x \cdot \frac{1}{\binom{m+1}{2}} = \frac{1}{\binom{m+1}{2}} \cdot \frac{m(m+1)(2m+1)}{2m+1} \\
    f_Y(y) &= \sum_{x=y}^{m} f_{X,Y}(x, y) = \frac{y}{\binom{m+1}{2}} \\
    f_{X/Y}(x/y) &= \frac{f_{X,Y}(x, y)}{f_Y(y)} = \frac{1}{\binom{m+1}{2}} \cdot \frac{1}{y} \\
    E[X/Y = y] &= \sum_{x=y}^{m} x \cdot \frac{1}{y(2m+1)} \\
    &= \frac{1}{m-y+1} \left[ \frac{m-y+1}{2} \right].
\end{aligned}
\]

Now, \( E[E(X/Y)] = E\left[ \frac{m+y}{2} \right] \).

\[
\begin{aligned}
    &= \frac{m}{2} + \frac{1}{2} E(Y) \\
    &= \frac{m}{2} + \frac{1}{2} \left\{ \sum_{y=1}^{m} y \cdot \frac{1}{\binom{m+1}{2}} \right\} \\
    &= \frac{m}{2} + \frac{1}{2} \cdot \frac{m(m+1)}{2}. \quad (\ast)
\end{aligned}
\]
**Joint MGF:**

- **Discrete case:** Let $(X_1, X_2)$ be a two-dimensional random variable with PMF $f_{X_1X_2}(x_1, x_2)$; the joint MGF of $(X_1, X_2)$ denoted by $M(t_1, t_2)$ is defined by

$$M(t_1, t_2) = \sum_{x_1, x_2} e^{t_1 x_1 + t_2 x_2} f_{X_1X_2}(x_1, x_2),$$

provided the expectation exists for all $(t_1, t_2)$ such that $|t_1|, |t_2| < \alpha_i$ for $i = 1, 2$.

- **Continuous case:** Let $(X_1, X_2)$ be a two-dimensional absolutely continuous random variable with joint PDF $f_{X_1X_2}(x_1, x_2)$, the joint MGF of $(X_1, X_2)$ denoted by $M(t_1, t_2)$ is defined by

$$M(t_1, t_2) = \int \int e^{t_1 x_1 + t_2 x_2} f(x_1, x_2) \, dx_1 \, dx_2,$$

provided the expectation exists for all $(t_1, t_2)$ such that $|t_1|, |t_2| < \alpha_i$ for $i = 1, 2$.

**Calculation of Moments:**

$$M(t_1, t_2) = \sum_{p=0}^\infty \sum_{q=0}^\infty t_1^p t_2^q E(X_1^p X_2^q) = \prod_{j=1}^{\infty} E(X_1^j) E(X_2^j) = E(Y) \times E(Y),$$

where $Y = X_1 X_2$.

- $\mu_{p,q}$ is the coefficient of $t_1^p t_2^q$ in $M(t_1, t_2)$.

$$M(t_1, t_2) = \left. \frac{d^2 M}{dt_1 dt_2} \right|_{t_1=0, t_2=0} = E(X_1^2 X_2^2),$$

$$\frac{d^2 M}{dt_1^2} \bigg|_{t_1=0} = E(X_1^2),$$

$$\frac{d^2 M}{dt_1 dt_2} \bigg|_{t_1=0, t_2=0} = E(X_1 X_2).$$
Example:

If \( X, Y \) has PMF \( f_{XY}(x,y) = \frac{e^{-2}}{x! (y-x)!}, x=0,1, \ldots \), \( y=0,1, \ldots, \infty \).

Find \( M(t_1, t_2) \).

\[
M(t_1, t_2) = \sum_{x=0}^{\infty} \sum_{y=0}^{\infty} e^{t_1 x + t_2 y} \frac{e^{-2}}{x! (y-x)!} \\
= \sum_{x=0}^{\infty} \frac{1}{x!} e^{-2} e^{t_2 y} \sum_{y=0}^{\infty} \frac{e^{t_1 x}}{y!} \\
= \frac{2}{y!} \left( e^{t_2} (1+e^{t_1}) \right)^y e^{-2} \\
= e^{-2} e^{t_2} (1+e^{t_1})
\]

\[\Rightarrow \begin{array}{c} \text{i.i.d. } N(0,1) \text{, find joint mgf of } Y_1 = X_1 + X_2, \text{ hence } Y_1, Y_2 \text{ are uncorrelated but not independent. } Y_2 = X_1 + X_2. \end{array}\]

\[
M(t_1, t_2) = E\left[ e^{t_1 Y_1 + t_2 Y_2} \right] \\
= E\left[ e^{t_1 X_1 + t_1 X_2 + t_2 X_1 + t_2 X_2} \right] \\
= e^{E(t_1 + t_2 X_1) + E(t_2 + t_2 X_2)} \text{ as } X_1, X_2 \text{ are i.i.d.}
\]

\[X \sim N(0,1)\]

\[
E\left[ e^{t_1 X_1 + t_2 X_1} \right] \\
= \int_{-\infty}^{\infty} e^{\frac{1}{2} t_1 x^2 - \frac{t_1}{2} x^2 (1-2t_2) - 2t_1 x_1} \, dx_1 \\
= \frac{1}{\sqrt{2\pi}} \left( \frac{1}{1-2t_2} \right)^{1/2} \left( \frac{1}{1-2t_2} \right)^{1/2} \int_{\frac{-t_1}{\sqrt{1-2t_2}}}^{\frac{t_1}{\sqrt{1-2t_2}}} \left[ \cdot \cdot \right] \, dt_1 \\
= e^{t_1 \sqrt{1-2t_2}} \left( e^{-t_1} \right) \text{ for } 0 < t_1 < \frac{1}{2}
\]

\( M(t_1, t_2) = \frac{1}{1-2t_2} \cdot e^{t_1 \sqrt{1-2t_2}} \).

Putting \( t_2 = 0, \) the marginal MGF of \( Y_1, \)

\( M_{Y_1}(t_1) = e^{t_1} \).

Putting \( t_1 = 0, \) the marginal MGF of \( Y_2, \)

\( M_{Y_2}(t_2) = \frac{1}{1-2t_2} \).

\( M(t_1, t_2) \neq M_{Y_1}(t_1) M_{Y_2}(t_2) \rightarrow Y_1 \text{ and } Y_2 \text{ are not independent.} \)

\( E(Y_1) = 0, \quad E(Y_2) = E(X_1) + E(X_2) = 2 \)

\( E(Y_1 Y_2) = E\left[ (X_1 + X_2) (X_1' + X_2') \right] = 0 \)

\( \therefore \text{cov}(Y_1, Y_2) = 0 \rightarrow Y_1 \text{ and } Y_2 \text{ are uncorrelated.} \)
Bivariate Normal Distribution: An absolutely continuous random vector \((X_1, X_2)\) is said to follow a bivariate normal distribution with parameters \((\mu_1, \mu_2, \Sigma)\) if the joint PDF of \((X_1, X_2)\) is

\[
f(x_1, x_2) = \frac{1}{\sqrt{2\pi \Sigma}} e^{-\frac{1}{2} (x_1 - \mu_1, x_2 - \mu_2) \Sigma^{-1} (x_1 - \mu_1, x_2 - \mu_2)}
\]

where \(\mu_1, \mu_2, \Sigma \in \mathbb{R}\) and \(|\Sigma| < 1\).

We can write \((X_1, X_2) \sim N(\mu_1, \mu_2, \Sigma)\).

The marginal PDF of \(X_1\) is

\[
f_{X_1}(x_1) = \int_{\mathbb{R}} f(x_1, x_2) \, dx_2 = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \Sigma}} e^{-\frac{1}{2} (x_1 - \mu_1)^2 \Sigma^{-1}} \, dx_2
\]

where

\[
f_{X_1}(x_1) = \frac{1}{\sqrt{2\pi \Sigma}} e^{-\frac{1}{2} (x_1 - \mu_1)^2 \Sigma^{-1}}, \quad -\infty < x_1 < \infty
\]

Note: The integrand is in the form of a normal PDF.

Similarly, it can be shown that

\[
f_{X_2}(x_2) = \frac{1}{\sqrt{2\pi \Sigma}} e^{-\frac{1}{2} (x_2 - \mu_2)^2 \Sigma^{-1}}, \quad -\infty < x_2 < \infty
\]

Conditional Distribution of \(X_2 / X_1 = \alpha_1; \quad \alpha_1 \neq 0\)

\[
f_{X_2 / X_1}(\alpha_2 / x_1) = \int_{\mathbb{R}} f_{X_2 / X_1}(\alpha_2 / x_1) \, dx_1 = \frac{1}{\sqrt{2\pi \Sigma(1 - \rho^2)}} \exp\left(-\frac{1}{2} \left(\frac{x_2 - \mu_2}{\sqrt{1 - \rho^2}}\right)^2\right)
\]
\[
\begin{align*}
\frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2} \frac{1}{\sigma_1^2(1-\rho^2)} (\tilde{x}_2 - \mu_2 - \rho \frac{\sigma_2}{\sigma_1} (\tilde{x}_1 - \mu_1))^2} \\
= \frac{1}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{1}{2} \frac{1}{\sigma_2^2(1-\rho^2)} (\tilde{x}_2 - \mu_2 - \rho \frac{\sigma_2}{\sigma_1} (\tilde{x}_1 - \mu_1))^2} \\
\end{align*}
\]

\[
\begin{align*}
\frac{x_2}{x_1} = \frac{x_1}{x_2} \sim N \left[ \frac{\mu_2 + \rho \frac{\sigma_1}{\sigma_2} (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2 + \rho^2 \sigma_2^2}{\sigma_1^2}}, \frac{\sigma_2(1-\rho^2)}{\sigma_2}} \right] \\
\frac{x_1}{x_2} = \frac{x_2}{x_1} \sim N \left[ \frac{\mu_1 + \rho \frac{\sigma_1}{\sigma_2} (\mu_2 - \mu_1)}{\sqrt{\frac{\sigma_1^2 + \rho^2 \sigma_2^2}{\sigma_1^2}}, \frac{\sigma_2(1-\rho^2)}{\sigma_2}} \right] \\
\end{align*}
\]

Moments:
\[
\begin{align*}
E(x_1) &= \mu_1 \\
V(x_1) &= \sigma_1^2 \\
E(x_2) &= \mu_2 \\
V(x_2) &= \sigma_2^2 \\
E(x_1 x_2) &= E \left[ E \left( \frac{x_2 x_1}{x_1} \right) \right] = E \left[ x_1 E \left( \frac{x_2}{x_1} \right) \right] \\
&= E \left[ x_1 \left( \mu_2 + \rho \frac{\sigma_1}{\sigma_2} (\mu_1 - \mu_2) \right) \right] \\
&= \mu_1 \mu_2 + \rho \frac{\sigma_1}{\sigma_2} \left[ E(x_1) - \mu_1 \right] \\
\end{align*}
\]

\[ p = \frac{E(x_1 x_2) - \mu_1 \mu_2}{\sigma_1 \sigma_2} \]

Problem: If \( (x_1, x_2) \sim N \left( \mu_1, \mu_2, \sigma_1, \sigma_2, \rho \right) \), \( x_1 \) and \( x_2 \) are independent iff \( \rho = 0 \).

Answer:
\[
\begin{align*}
\begin{pmatrix} x_1 \ x_2 \end{pmatrix} \\
\end{pmatrix} \sim \frac{1}{2\pi \sigma_1 \sigma_2} e^{-\frac{1}{2} \left( \frac{(x_1 - \mu_1)^2}{\sigma_1^2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} - \rho \frac{\sigma_1}{\sigma_2} (x_1 - \mu_1)(x_2 - \mu_2) \right)} \\
\end{align*}
\]

Now, if \( x_1 \) and \( x_2 \) are independent,
\[
\begin{align*}
E(x_1 x_2) &= E(x_1) E(x_2) \\
\text{Cov}(x_1, x_2) &= 0 \Rightarrow \rho = 0.
\end{align*}
\]
MGF of Bivariate Normal Distribution:

Joint MGF of \((X_1, X_2)\) is

\[
M(t_1, t_2) = \mathbb{E} \left[ e^{t_1 X_1 + t_2 X_2} \right]
\]

\[
= \mathbb{E} \left[ e^{t_1 X_1} \mathbb{E} \left[ e^{t_2 X_2} / X_1 \right] \right]
\]

\[
= \mathbb{E} \left[ e^{t_1 X_1} e^{t_2 \frac{X_2 - \theta_1}{\sigma_1}} \left( X_1 - \theta_1 \right) + \frac{t_2 \sigma_2^2 (1 - \rho^2)}{2} \right]
\]

\[
= e^{\mu_2 t_2 + \frac{t_2 \sigma_2^2}{2}} \left( 1 - \frac{t_2 \rho \sigma_2}{\sigma_1} \right) - \mu_1 t_2 e^{\frac{t_2 \sigma_2^2}{2}} \mathbb{E} \left[ e^{t_1 X_1} + t_2 e^{\frac{t_2 \sigma_2^2}{2}} \right]
\]

\[
= e^{\mu_2 t_2 + \frac{t_2 \sigma_2^2}{2}} \left( 1 - \frac{t_2 \rho \sigma_2}{\sigma_1} \right) - \mu_1 t_2 e^{\frac{t_2 \sigma_2^2}{2}} \mathbb{E} \left[ e^{t_1 X_1} + t_2 e^{\frac{t_2 \sigma_2^2}{2}} \right]
\]

\[
= e^{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2} \left[ \theta_1 e^{\frac{t_2 \sigma_2^2}{2}} + 2 \sigma_1 \sigma_2 \rho e^{\frac{t_2 \sigma_2^2}{2}} \right]}
\]

\[
M(t_1, t_2) = e^{\mu_1 t_1 + \mu_2 t_2 + \frac{1}{2} \left[ \theta_1 e^{\frac{t_2 \sigma_2^2}{2}} + 2 \sigma_1 \sigma_2 \rho e^{\frac{t_2 \sigma_2^2}{2}} \right]}
\]

\[\text{[Expectation exists if } t_1, t_2]\]

Problem 2. If \((X, Y) \sim BN(0, 0, \sigma_1^2, \sigma_2^2, \rho)\)

Find the correlation coefficient between \(X, Y\).

\[\begin{align*}
\text{Ans:} & \quad \mathbb{E} (X) = \sigma_1^2 + \mu_1^2 = \sigma_1^2 \\
E (X^2) &= \sigma_2^2, \\
E (X^2 Y^2) &= \mathbb{E} \left[ X^2 \mathbb{E} \left( Y^2 / X \right) \right] \\
&= \mathbb{E} \left[ X^2 \left( \sigma_1^2 (1 - \rho^2) + \sigma_2^2 \frac{\sigma_1^2}{\sigma_2^2} X_1^2 \right) \right] \\
&= \sigma_2^2 (1 - \rho^2) \sigma_1^2 + \frac{\sigma_2^4 \sigma_1^2}{\sigma_2^2} \\
&= \sigma_2^2 (1 - \rho^2) \sigma_1^2 + 3 \sigma_2^4 \\
\end{align*}\]

\[\begin{align*}
\gamma (X, Y) &= 3 \sigma_1^4 - \sigma_1^4 = 2 \sigma_1^4 \\
\rho (X, Y) &= 2 \sigma_2^4 \\
\rho (X, Y) &= \frac{\sigma_1^2 \sigma_2^2 (1 - \rho^2) - \sigma_1^2 \sigma_2^2 + 3 \sigma_2^4 \sigma_1^4}{2 \sigma_1^2 \sigma_2^2} \]
\[= \rho^2 \]
Problem 8: Give an example of a joint distribution of a 2-dimensional random vector \((X_1, X_2)\)

\(X_1^2, X_2^2\) are independent but \(X_1\) and \(X_2\) are not.

**ANS:** Consider the joint PDF

\[
f(x_1, x_2) = \frac{1}{4} (1 + x_1 x_2), \quad |x_1| \leq 1, |x_2| \leq 1
\]

\[
f(x_i) = \frac{1}{4} \int (1 + x_1 x_2) \, dx_2 = \frac{1}{4} [1 + 1 + 0] = \frac{1}{2}
\]

Let \(U = x_1^2, \quad V = x_2^2\).

The joint distribution function of \(U\) and \(V\) is

\[
F(u, v) = P[U \leq u, \quad V \leq v]
\]

\[
= \int_{\sqrt{u}}^{1} \int_{\sqrt{v}}^{1} \frac{1}{4} (1 + x_1 x_2) \, dx_1 \, dx_2
\]

\[
= \frac{1}{4} \left[ 2 \left( \sqrt{u}, \sqrt{v} \right)^2 \right]
\]

\[
= \frac{u}{4} \left( \sqrt{u}, \sqrt{v} \right)
\]

\[
P[U \leq u] = \frac{1}{2} \int_{\sqrt{u}}^{1} du = \sqrt{u}
\]

\[
P[V \leq v] = \frac{1}{2} \int_{\sqrt{v}}^{1} dv = \sqrt{v}
\]

\[
F(u, v) = F(u)F(v)
\]

\(X_1^2\) and \(X_2^2\) are independent but \(X_1\) and \(X_2\) are not.

**Note:**

\((X_1, X_2) \sim BN(0, 0, 1, 1, \rho)\)

\(X_1/X_2 \sim N(\sqrt{\rho^2}, \sqrt{1-\rho^2})\).

\[
E[X_1 X_2^{2\rho}] = E[E[X_1 X_2^{2\rho} \mid X_2]]
\]

\[
= E[X_2^{2\rho} E[X_1 \mid X_2]]
\]

\[
= E[X_2^{2\rho}] \rho = \rho X_0
\]
Problem 4. If \((X_1, X_2) \sim N_2(0, 0, 1, 1, \rho)\), then show that:

i) \(P[X_1 > 0, X_2 > 0] = \frac{1}{4} + \frac{1}{2\pi} \sin^{-1}(\rho)\)

ii) \(P[X_1 < 0, X_2 < 0] = \frac{1}{4} - \frac{1}{2\pi} \sin^{-1}(\rho)\)

iii) \(P[X_1 > 0, X_2 < 0] = \frac{1}{4} - \frac{1}{2\pi} \sin^{-1}(\rho)\)

iv) \(P[X_1 < 0, X_2 > 0] = \frac{1}{4} - \frac{1}{2\pi} \sin^{-1}(\rho)\)

Solution:

The joint PDF of \((X_1, X_2)\) is

\[
 f_{X_1 X_2}(x_1, x_2) = \frac{1}{2\pi \sqrt{1 - \rho^2}} e^{-\frac{1}{2(1-\rho^2)} (x_1^2 + x_2^2 - 2\rho x_1 x_2)}
\]

Let us consider the transformation,

\[
\begin{align*}
&\alpha_1 = r \cos \theta, \quad r > 0 \\
&\alpha_2 = r \sin \theta, \quad 0 < \theta < \pi/2
\end{align*}
\]

\(\alpha_1 + \alpha_2 = r^2\)

Jacobian of the transformation is,

\[
J = \begin{vmatrix}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{vmatrix} = r
\]

\(\therefore |J| = r\) \(\because r > 0\)

\[
P[X_1 > 0, X_2 > 0] = \frac{1}{2\pi \sqrt{1 - \rho^2}} \int_0^{\pi/2} \int_0^{\infty} e^{-\frac{1}{2(1-\rho^2)} (r^2 \cos^2 \theta + r^2 \sin^2 \theta - 2\rho r^2 \cos \theta \sin \theta)} r \ dr \ d\theta
\]

\[
= \frac{1}{2\pi \sqrt{1 - \rho^2}} \int_0^{\pi/2} \int_0^{\infty} e^{-\frac{r^2}{2(1-\rho^2)} (1 - \rho \sin 2\theta)} r \ dr \ d\theta
\]

Put, \(\frac{r^2}{2(1-\rho^2)} = z\)

\[
\therefore \text{Area} = \frac{1}{1 - \rho \sin 2\theta} \cdot \frac{1}{2\pi \sqrt{1 - \rho^2}} \int_0^{\pi/2} \int_0^{\infty} e^{-z} \ dz \ d\theta
\]
\[
\begin{align*}
\therefore P[X_1 > 0, X_2 > 0] &= \int_0^\infty \int_0^\infty \frac{1}{2\pi (1 - p^2)} e^{-\frac{r^2}{2(1 - p^2)}} r dr d\theta \\
&= \frac{1}{2\pi} \int_0^\pi \int_0^{\sqrt{1 - p^2}} \frac{1}{1 + \tan^2 \theta} \frac{1}{2\pi} \frac{1}{1 - p^2} (1 - p^2) \tan^{-1} \left( \frac{\sqrt{1 - p^2}}{p} \right) \, dy \\
&= \frac{1}{2\pi} \left( \frac{\pi}{2} + \tan^{-1} \left( \frac{\sqrt{1 - p^2}}{p} \right) \right) \\
&= \frac{1}{4} + \frac{1}{2\pi} \tan^{-1} (p) \\
\end{align*}
\]

Let \( A = \{ X_1 > 0 \} \) and \( B = \{ X_2 > 0 \} \)

i) \( P[X_1 > 0, X_2 > 0] = P[A \cap B] = P(A) - P(A \cap B) \)

\( = P(A) - P(B) + P[A^c \cap B^c] \)

\( = \frac{1}{2} - \frac{1}{2} + P[X_1 < 0, X_2 < 0] \)

\( = \frac{1}{2} + \frac{1}{2\pi} \tan^{-1} (p) \)

ii) \( P[X_1 > 0, X_2 < 0] = P[X_1 > 0] - P[X_1 > 0, X_2 > 0] \)

\( = \frac{1}{2} - \frac{1}{4} - \frac{1}{2\pi} \tan^{-1} (p) \)

\( = \frac{1}{4} - \frac{1}{2\pi} \tan^{-1} (p) \)

iii) \( P[X_1 < 0, X_2 > 0] = P[X_1 < 0, X_2 < 0] \)

\( = \frac{1}{4} - \frac{1}{2\pi} \tan^{-1} (p) \)

\[ \text{Continuation:} \quad \text{If} \quad q = P[X_1 < 0, X_2 < 0] \quad \text{then} \quad \pi q = \cos (\pi q) \]

\[ \text{Ans:} \quad q = P[X_1 < 0, X_2 < 0] = P[X_1 < 0, X_2 < 0] + P[X_1 > 0, X_2 < 0] \]

\( = \frac{1}{2} - \frac{1}{2\pi} \tan^{-1} (p) \)

on, \( P = \sin \left( \frac{\pi}{2} - \pi q \right) = \cos (\pi q) \), where,

\( (x, x_2) \sim N_2 (0, 0, 1, 1, \rho) \)
\[ V(Y) = E[V(Y/X)] + V[E(Y/X)] \\
= E[\Sigma^2_y (1-\rho^2)] + \mu^2 + \rho \frac{\Sigma_y}{\sigma_x} (\mu - \mu_x) \\
= \Sigma^2_y (1-\rho^2) \\
E(\text{XY}) = E[E(\text{XY}/X)] = E[X \cdot \mu^2 + \rho \frac{\Sigma_y}{\sigma_x} (\mu - \mu_x) E(X)] \\
= \mu^2 E(X) + \rho \frac{\Sigma_y}{\sigma_x} \mu_x E(X) - \rho \frac{\Sigma_y}{\sigma_x} \mu E(X) \\
\text{Cov}(X,Y) = \Sigma^2_Y E(X) + \rho \frac{\Sigma_y}{\sigma_x} E(X^2) - \rho \frac{\Sigma_y}{\sigma_x} \mu_x E(X) - \mu^2 E(X) \\
= \rho \frac{\Sigma_y}{\sigma_x} V(X). \\
\therefore \frac{\rho_{xy}}{\sigma_x \sigma_y} = \frac{\rho \frac{\Sigma_y}{\sigma_x} V(x)}{\sqrt{\Sigma(Y)^2 (1-\rho^2)^2 + \rho^2 \Sigma_Y^2} V(x)} \\
= \frac{1}{\sqrt{1 + \frac{\Sigma_X^2 (1-\rho^2)}{\rho^2}}} \\
\star \text{Problem 5.} \frac{\text{Let } f_1(x,y), f_2(x,y) \text{ be two bivariate normal density functions with zero means, unit variances and correlation coefficients } \rho_1, \rho_2 \text{ respectively } (\rho_1 \neq \rho_2), \text{ show that the density function }} \\
\frac{f(x,y) = \frac{1}{2} f_1(x,y) + \frac{1}{2} f_2(x,y)}{\text{is not normal. But its density functions are normal.}} \\
\text{Solution } \\
\frac{f_1(x,y)}{\frac{\Sigma_1^2}{\Sigma^2} \frac{1}{\sqrt{1-\rho_1^2}}} e^{-\frac{1}{2(1-\rho_1^2)} (\alpha^2 + \gamma^2 - 2\rho_1 \alpha \gamma)} \\
\frac{f_2(x,y)}{\frac{\Sigma_2^2}{\Sigma^2} \frac{1}{\sqrt{1-\rho_2^2}}} e^{-\frac{1}{2(1-\rho_2^2)} (\alpha^2 + \gamma^2 - 2\rho_2 \alpha \gamma)} \\
\frac{f(x,y)}{\frac{\Sigma_1^2}{\Sigma^2} \frac{1}{\sqrt{1-\rho_1^2}}} e^{-\frac{1}{2(1-\rho_1^2)} (\alpha^2 + \gamma^2 - 2\rho_1 \alpha \gamma)} + \frac{1}{\sqrt{1-\rho_2^2}} e^{-\frac{1}{2(1-\rho_2^2)} (\alpha^2 + \gamma^2 - 2\rho_2 \alpha \gamma)} \\
\therefore \frac{f(x)}{\frac{\Sigma_1^2}{\Sigma^2} \frac{1}{\sqrt{1-\rho_1^2}}} e^{-\frac{1}{2(1-\rho_1^2)} (\alpha^2 + \gamma^2 - 2\rho_1 \alpha \gamma)} \\
\therefore \frac{f(y)}{\frac{\Sigma_2^2}{\Sigma^2} \frac{1}{\sqrt{1-\rho_2^2}}} e^{-\frac{1}{2(1-\rho_2^2)} (\alpha^2 + \gamma^2 - 2\rho_2 \alpha \gamma)} \\
\therefore X \sim N(0,1) \\
\text{Similarly, } Y \sim N(0,1).
Mode: \[ f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2} \left( \frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \right)^2} \]

Solve \[ f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2} \left( \frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \right)^2} \]

\[ \therefore \] if \( f \) is maximum when \( g \) is maximum.

\[ g(x_1, x_2) = \frac{1}{\sqrt{1-\rho^2}} \left( \frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \right)^2 \]

\[ = \left( \frac{x_1 - \mu_1}{\sigma_1} \right)^2 + \left( \frac{x_2 - \mu_2}{\sigma_2} \right)^2 - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \]

\[ \therefore \] if \( \frac{x_1 - \mu_1}{\sigma_1} = 0 \Rightarrow x_1 = \mu_1 \]

\[ \therefore x_2 = \mu_2 \]

\[ \therefore \] mode is at \( (\mu_1, \mu_2) \).

Problem: If \( x_1 \sim N(\mu_1, \sigma_1^2) \) & \( x_2 \sim N(\mu_2, \sigma_2^2) \), s.t. the joint distr. of \( (x_1, x_2) \) is bivariate normal, obtain the parameters of this distribution.

Solve \[ f(x_1, x_2) = f(x_1) \cdot f(x_2) \]

\[ = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2} \left( \frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \right)^2} \times g(x_1, x_2) \]

\[ = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{1}{2} \left( \frac{x_1 - \mu_1}{\sigma_1} + \frac{x_2 - \mu_2}{\sigma_2} - 2\rho \frac{(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} \right)^2} \]

\[ \therefore \] \( g(x_1, x_2) = (x_1 - \mu_1)^2 + (x_2 - \mu_1)^2 \]

\[ = (x_1 - \mu_1)^2 + (x_2 - \mu_2)^2 - 2(x_1 - \mu_1)(x_2 - \mu_2) \]

Hence, the dispersion matrix is p.d. \[ \det \begin{bmatrix} 2 & -\rho \frac{\sigma_1}{\sigma_2} \\ -\frac{\sigma_1}{\sigma_2} & 1 \end{bmatrix} = \frac{\sigma_1^2}{\sigma_2^2} > 0 \]

So, the joint distribution of \( (X_1, X_2) \) is bivariate normal distribution.

\[ \mu = \text{E}(X_2) = \text{E}[\text{E}(X_2/X_1)] = \mu_2 \]

\[ \sigma^2 = \text{Var}(X_2) + \text{Var}(\text{E}(X_2/X_1)) \]

\[ = \text{E}(\sigma^2) + \text{Var}(\mu_1) \]

\[ = 2\sigma_1^2 \]

\[ \sigma^2 = \text{E}[\text{E}(X_1X_2/X_1)] = \text{E}[X_1 \cdot \text{E}(X_2/X_1)] = \text{E}(X_1^2) \]

\[ = \frac{(\sigma_1^2)}{\sqrt{1-\rho^2}} \]
A Problem. Show that for the dist. $N_2(0,0,1,1,\rho)$, the moments obey the recurrence relation

$$\mu_{n+s} = (n+s-1)\rho\mu_{n-1}s-1 + (n-1)(s-1)(1-\rho^2)\mu_{n-2}s-2.$$ 

Hence, on otherwise, show that

$$\mu_{n+s} = 0 \quad \text{if} \quad n+s \text{ is odd};$$

$$\mu_{13} = \mu_{81} = 8\rho; \quad \mu_{22} = 1 + \rho^2.$$ 

Thus, $\mu_{n, s} = E(X^nY^s)$

$$= E(X^nY^{s-1}(Y-\rho X + \rho X))$$

$$= E\left[ X^n Y^{s-1}(Y-\rho X) \right] + P E(X^{n+1}Y^{s-1})$$

$$= E\left[ X^n Y^{s-1}(Y-\rho X) \right] + P \mu_{n+1}s-1.$$ 

Note,

$$E(X^nY^{s-1}(Y-\rho X)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^n y^{s-1}(y-\rho x) \frac{1}{2\pi \sqrt{1-\rho^2}} \, dx \, dy.$$ 

$$= \int_{-\infty}^{\infty} \frac{x^n e^{-x^2/2}}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} y^{s-1}(y-\rho x) \frac{1}{\sqrt{2\pi \sqrt{1-\rho^2}}} \, dy \right\} \, dx$$

$$= \int_{-\infty}^{\infty} \frac{x^n e^{-x^2/2}}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} (s-1)y^{s-2}\frac{1}{\sqrt{2\pi \sqrt{1-\rho^2}}} \, dy \right\} \, dx$$

$$= (s-1)(1-\rho)\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi \sqrt{1-\rho^2}}} \, dy \, dx$$

$$= (s-1)(1-\rho) \mu_{n, s-2}.$$ 

Interchanging the roles of $n$ and $s$, it can be similarly shown that

$$\mu_{n,s} = (n-1)(1-\rho)\mu_{n-2,s-1} + P\mu_{n+1,s-1}.$$ 

Using (1),

$$\mu_{n,s} = (n-1)(1-\rho)\mu_{n-2,s-2} + P\mu_{n+1,s-1}.$$ 

Using (2),

$$\mu_{n+1,s-1} = \rho_{n} (1-\rho)\mu_{n-1,s-1} + P\mu_{n,s-1}.$$
Thus, $\mu_{n,s} = (1-p^z) \left( (1-p^y) \mu_{n-1, s-2} + p \mu_{n-1, s-1} \right)
+ p \left( (1-p^y) \mu_{n-1, s-1} + p \mu_{n, s} \right)
\Rightarrow \mu_{n,s} = (3-1)(1-p^y) \mu_{n-1, s-2} + p(\mu_{n-1, s-1} + \mu_{n, s})$

if $n+s$ is odd, \( s \neq \).

Set, \( n > s \)

Now, \( \mu_{n,s} = c_1 \mu_{n-2, s-2} + c_2 \mu_{n-1, s-1} \)
\[= c_1 \left[ c_3 \mu_{n-4, s-4} + c_4 \mu_{n-3, s-3} \right]
+ c_2 \left[ c_5 \mu_{n-2, s-3} + c_6 \mu_{n-2, s-2} \right] \]
\[= k_1 \mu_{n-4, s-4} + k_2 \mu_{n-3, s-3} \]
\[\vdots\]
\[= k_8 \mu_{n-8, s} + k_9 \mu_{n-7, s+1} \]

\( \mu_{n-8, s} = 0 \) since \( n-s \) is odd

\( \mu_{n-8, s+1} = 0 \) since \( n-s+2=0 \) odd

Similarly, it can be shown that

if \( n < s \), \( \mu_{n, s} = 0 \) if \( n+s \) is odd.

Regression: Consider two variables \( X \) and \( Y \) where \( Y \) is the
variable. Our problem is to predict \( Y \), given \( X \) is known.

Let \( \eta_x = E[Y|X=x] \) provided it exists. The regression
curve of \( Y \) on \( X \) is defined by the locus of the point \( (x, \eta_x) \)
average of \( Y \) on \( X \), is defined by the locus of the point \( (x, \eta_x) \)
where \( \eta_x \) is any function of \( Y \) based on \( X \).

\[ E[Y-g(x)]^2 = E[(Y-\eta_x) - (g(x)-\eta_x)]^2 \]
\[ = E[(Y-\eta_x)^2] + E[(g(x)-\eta_x)^2] - 2E[(Y-\eta_x)(g(x)-\eta_x)] \]
\[ = E[(Y-\eta_x)^2] + E[(g(x)-\eta_x)^2] - 2E[(Y-\eta_x)(g(x)-\eta_x)] \]
\[ = E[(Y-\eta_x)^2] + E[(g(x)-\eta_x)^2] \geq 0 \]

\( E[Y-g(x)]^2 \) is minimum when
\[ g(x) = \eta_x \]

where, \( \eta_x = \frac{E[(Y-\eta_x)(g(x)-\eta_x)]}{E[(g(x)-\eta_x)^2]} \)
\[ = E[(g(x)-\eta_x)(Y-\eta_x)] = 0 \]

\[ = E[(g(x)-\eta_x)(Y-\eta_x)(X)/x^2] \]
Correlation Coefficient between $Y$ and $\eta_x$:

$$\text{Cov}(Y, \eta_x) = E[(Y - E(Y))(\eta_x - E(\eta_x))]$$
$$= E[(Y - E(Y))(\eta_x - E(\eta_x))]$$
$$= E[(\eta_x - E(\eta_x))E[(Y - E(Y))/x]]$$
$$= E[\eta_x - E(\eta_x)]^2$$
$$= \rho_{\eta_x}$$

$$\rho_{Y, \eta_x} = \frac{\text{Cov}(Y, \eta_x)}{\sqrt{\text{Var}(Y)\text{Var}(\eta_x)}} = \frac{\text{Cov}(Y, \eta_x)}{\sqrt{\text{Var}(Y)\text{Var}(\eta_x)}}$$

$$\rho_{Y, \eta_x} ^ 2 = \frac{\text{Cov}^2(Y, \eta_x)}{\text{Var}(Y)\text{Var}(\eta_x)} = \frac{\rho_{\eta_x} ^ 2}{\text{Var}(Y)\text{Var}(\eta_x)}$$

$$\rho_{Y, g(x)} = \frac{\text{Cov}(Y, g(x))}{\sqrt{\text{Var}(Y)\text{Var}(g(x))}} = \frac{\rho_{\eta_x} \rho_{g(x)}}{\sqrt{\text{Var}(Y)\text{Var}(\eta_x)\text{Var}(g(x))}}$$

$$\rho_{Y, g(x)} ^ 2 = \frac{\rho_{\eta_x} ^ 2 \rho_{g(x)} ^ 2}{\text{Var}(Y)\text{Var}(\eta_x)\text{Var}(g(x))}$$

Note: The square of the maximum correlation attained is called the correlation ratio of $Y$ on $X$ and is usually denoted by

$$\eta_{XY}^2 = \rho_{Y,x}^2, \quad 0 \leq \eta_{XY}^2 \leq 1 \quad [\text{To show}]$$

$$\eta_{XY}^2 = \frac{\text{Var}(\eta_x)}{\text{Var}(Y)}$$

$$\text{Var}(Y) = E[(Y - E(Y))^2] = E[\text{Var}(Y/X)] + E[E(\text{Var}(Y/X))] \geq 0$$

$$\frac{\text{Var}(\eta_x)}{\text{Var}(Y)} \leq 1 \Rightarrow \eta_{XY}^2 \leq 1$$
Problem 1. Let $X$ and $Y$ be two jointly distributed continuous random variables. Consider their joint PDF,

$$f_{XY}(x, y) = \frac{1}{2\pi} \exp \left[ -\frac{1}{2} \left( \frac{x^2 - 2pxy + y^2}{1-p^2} \right) \right], x \in \mathbb{R}, y \in \mathbb{R}$$

i) Find the marginal PDF of $X$.

ii) Find the conditional PDF of $Y$ for given $X = x$.

Solution:

i) $f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dy$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2(1-p^2)} \left( x^2 - 2pxy + y^2 \right) \right] \, dy$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2(1-p^2)} \left( y^2 - (1-p^2)x^2 \right) \right] \, dy$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{(y-px)^2}{2(1-p^2)}} \, dy$$

Since,

$$\int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2a^2} (y - \mu)^2 \right] \, dy = 1$$

Here, $\mu = px$ and $\sigma^2 = (1-p^2)$

$$\int_{-\infty}^{\infty} \exp \left[ -\frac{1}{2(1-p^2)} (y-px)^2 \right] \, dy = \sqrt{2\pi(1-p^2)}$$

$$f_X(x) = \frac{1}{\sqrt{2\pi(1-p^2)}} e^{-\frac{(y-px)^2}{2(1-p^2)}}, \quad x \in \mathbb{R}$$

ii) Conditional PDF of $Y$ for given $X = x$ is $-\frac{(y-px)^2}{2(1-p^2)} \quad -\infty < y < \infty$

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)} = \frac{1}{\sqrt{2\pi(1-p^2)}} e^{-\frac{(y-px)^2}{2(1-p^2)}}$$

i.e., $Y/X = x \sim N \left( px, \frac{(1-p^2)}{2(1-p^2)} \right)$, $-\infty < y < \infty$

$$E[Y/X = x] = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) \, dy = px$$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y) \, dy.$$
Problem 2. \( f(x, y) = \begin{cases} \infty & \text{if } 0 < x < y \text{ and } 0 < y < \infty \\ x & \text{if } 0 < y < x \text{ and } 0 < x < \infty \end{cases} \)

i) Find marginal PDF of \( X \).
ii) Find conditional PDF of \( Y \) for given \( X = x \) and also \( E(Y/X = x) \).

Solution: The marginal PDF of \( X \) is
\[
 f_X(x) = \int_0^\infty f_{X,Y}(x,y) \, dy
 = \int_0^x f_{X,Y}(x,y) \, dy + \int_x^\infty f_{X,Y}(x,y) \, dy
 = \int_0^x e^{-x}(1-e^{-y}) \, dy + \int_x^\infty e^{-y}(1-e^{-x}) \, dy
 = e^{-x} [x + e^{-x} - 1] + e^{-x} (1 - e^{-x})
 = xe^{-x}, \quad 0 < x < \infty
\]

for some \( \alpha \), where \( x > 0 \).
\[
 f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}
 = \begin{cases} \frac{e^{-y}(1-e^{-x})}{xe^{-x}} = \frac{e^{-y} - e^{-x}}{x}, & \text{if } 0 < y < x \\ \frac{1-e^{-y}}{x}, & \text{if } 0 < x < y \end{cases}
\]

\[
 E[Y/X = x] = \int_0^\infty y f_{Y|X}(y|x) \, dy
 = \int_0^x y \frac{e^{-y} - e^{-x}}{x} \, dy + \int_x^\infty y \frac{1-e^{-y}}{x} \, dy
 = \frac{1}{x} [x + e^{-x} - 1] + \frac{1}{x} [1 - e^{-x}]
 = 1.
\]
Problem 3. Let $X$ and $Y$ be two jointly distributed continuous random variables with PDF

$$f_{XY}(x,y) = \frac{1}{8} (y - x^2) e^{-y} \text{ when, } 0 < y < x, |x| < y,
$$
i.e. $-y < x < y$.

Solution:

$$f_X(x) = \begin{cases} 
\int_0^x f_{XY}(x,y) \, dy & \text{for } x > 0 \\
\int_{-x}^0 f_{XY}(x,y) \, dy & \text{for } x < 0 
\end{cases}$$

When $x > 0$,

$$f_X(x) = \frac{1}{8} \int_x^\infty (y - x^2) e^{-y} \, dy$$

$$= \frac{1}{8} \left[ -y e^{-y} \right]_x^\infty + \frac{1}{8} \int_x^\infty y e^{-y} \, dy$$

$$= \frac{1}{8} \left[ -y e^{-y} \right]_x^\infty + \frac{1}{8} \left[ -e^{-y} \right]_x^\infty + \frac{1}{8} \left[ e^{-y} \right]_x^\infty$$

$$= \frac{1}{8} x e^{-x} + \frac{1}{4} e^{-x} + \frac{1}{4} e^{-x} - \frac{x^2}{8} e^{-x}$$

When $x < 0$,

$$f_X(x) = \frac{1}{8} \int_{-x}^0 (y - x^2) e^{-y} \, dy$$

$$= \frac{1}{8} \left[ y e^{-y} \right]_{-x}^0 + \frac{1}{8} \left[ -y e^{-y} \right]_{-x}^0$$

$$= \frac{1}{8} e^{-x} + \frac{1}{8} x e^{-x}$$
Example 1: Let \( X \) and \( Y \) be two jointly distributed \( \text{pdf} \):
\[
f_{XY}(x, y) = \begin{cases} \frac{x^2}{2}, & \text{if } 0 < x < 1, 0 < y < 2 \\ 0, & \text{otherwise} \end{cases}
\]

Find:

i) the marginal PDF of \( X \)

ii) Marginal PDF of \( Y \)

iii) Compute \( P[X > \frac{1}{2}], P[Y < X], P[Y < \frac{1}{2} / X < \frac{1}{2}] \)

Solution:

i) Marginal PDF of \( X \) is:
\[
f_X(x) = \int_0^2 f_{XY}(x, y) \, dy = \int_0^2 \left( \frac{x^2}{2} + \frac{2xy}{3} \right) \, dy
\]
\[
= 2x^2 + \frac{2x}{3}, \quad 0 < x < 1.
\]

ii) Marginal PDF of \( Y \) is:
\[
f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y) \, dx = \int_0^1 \left( \frac{x^2}{2} + \frac{2xy}{3} \right) \, dx
\]
\[
= \left( \frac{x^3}{3} + \frac{xy^2}{6} \right) \bigg|_0^1 = \frac{1}{3} + \frac{y^2}{6}
\]

iii) \( P[Y < X] = \int \int_{0 < y < x} f_{XY}(x, y) \, dx \, dy = \int_0^1 \int_0^x \frac{x^2}{2} \, dy \, dx
\]
\[
P[Y < \frac{1}{2} / X < \frac{1}{2}] = \frac{P[X < \frac{1}{2}, Y < \frac{1}{2}]}{P[X < \frac{1}{2}]}
\]
\[
= \frac{\int_{1/2}^{1} \int_0^{1/2} f_{XY}(x, y) \, dx \, dy}{\int_0^{1/2} \int_0^{1} f_{XY}(x, y) \, dx \, dy}
\]
\[
= \frac{\int_{1/2}^{1} \int_0^{1/2} \left( \frac{x^2}{2} + \frac{2xy}{3} \right) \, dx \, dy}{\int_0^{1/2} \int_0^{1} \left( \frac{x^2}{2} + \frac{2xy}{3} \right) \, dx \, dy}
\]
Let \( X \) and \( Y \) be two jointly distributed r.v.'s so the marginal PDF of \( X \) is
\[
    f_X(x) = \begin{cases} 
      1, & \text{if } -\frac{1}{2} < x < \frac{1}{2} \\
      0, & \text{otherwise}
    \end{cases}
\]
also let the conditional PDF of \( Y \) given \( X \) be given by
\[
    f_{Y\mid X}(y\mid x) = \begin{cases} 
      1, & \text{if } -x < y < x+1 \\
      0, & \text{otherwise}
    \end{cases}
\]
for \(-\frac{1}{2} < x < 0\) and by
\[
    f_{Y\mid X}(y\mid x) = \begin{cases} 
      1, & \text{if } -x < y < 1-x \\
      0, & \text{otherwise}
    \end{cases}
\]
for \(0 < x < \frac{1}{2}\).

Show that \( X \) and \( Y \) are uncorrelated. Are they independent?

**Solution:** The joint PDF of \( X \) and \( Y \) is
\[
    f_{X\mid Y}(x\mid y) = f_{Y\mid X}(y\mid x) f_X(x)
\]
Note that,
\[
    E(XY) = \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-x}^{x+1} xy f_{X\mid Y}(x\mid y) \, dy \, dx + \int_{0}^{\frac{1}{2}} \int_{-x}^{1-x} xy f_{X\mid Y}(x\mid y) \, dy \, dx
\]
\[
    = 0 + \frac{1}{2} - \frac{1}{2} = 0
\]
\[
    E(X) = \int_{-\frac{1}{2}}^{\frac{1}{2}} x f_X(x) \, dx = 0
\]
Hence, \( X \) and \( Y \) are uncorrelated.
**Problem 6.** Let X and Y have the circular normal distribution with mean \( (0, 0) \), i.e., \( X \& Y \sim N_2 \left( 0, 0, \sigma_x^2, \sigma_y^2, 0 \right) \). Consider a circle \( C \) and a square \( S \) of equal area both with a centre \( (0, 0) \).

Prove that, \( P \left[ \begin{pmatrix} X \& Y \end{pmatrix} \in C \right] > P \left[ \begin{pmatrix} X \& Y \end{pmatrix} \in S \right] \) \quad (2006)

**Solution:**

The joint PDF of \( X \& Y \) is given by

\[
    f(x, y) = \frac{1}{2\pi \sigma^2} e^{-\frac{1}{2\sigma^2} (x^2 + y^2)} \quad \forall x, y \in \mathbb{R} \quad \sigma > 0
\]

Let us consider a square \( S \) with vertices \((a, -a), (a, a), (-a, a), (-a, -a)\).

The area of the square = \( 4a^2 = S \)

Consider a circle \( C \) with radius \( r \), and the centre at \((0, 0)\).

Area of \( C \) = \( \pi r^2 \)

Hence, \( \pi r^2 = 4a^2 \) \quad [given]

\[
    \Rightarrow r = \frac{2a}{\sqrt{\pi}}
\]

Therefore, \( a < r < \sqrt{2}a \)

Now, \( P \left[ \begin{pmatrix} x \& y \end{pmatrix} \in S \right] = \int_{-a}^{a} \int_{-a}^{a} f(x, y) \, dx \, dy \)

\[
    = 4 \int_{0}^{a} \int_{0}^{a} f(x, y) \, dx \, dy \quad [By \ symmetry]
\]

\( P \left[ \begin{pmatrix} x \& y \end{pmatrix} \in C \right] = \int_{-a}^{a} \int_{-a}^{a} f(x, y) \, dx \, dy \)

Now, in the first quadrant,

\[
    P\left[ \begin{pmatrix} x \& y \end{pmatrix} \in C \right] - P\left[ \begin{pmatrix} x \& y \end{pmatrix} \in S \right] = \int_{a}^{\infty} \int_{a}^{\infty} f(x, y) \, dx \, dy - \int_{a}^{\infty} \int_{a}^{\infty} f(x, y) \, dx \, dy \quad [From \ the \ figure, \ cancelling \ the \ common \ region]
\]

\( A = \text{shaded region,} \)

\( B = \text{dotted region,} \)
Now if \((x, y) \in A\), then,
\[
\begin{align*}
& (x^2 + y^2) < r^2 \\
\Rightarrow & -\frac{(x^2 + y^2)}{2\sigma^2} > -\frac{r^2}{2\sigma^2} \\
\Rightarrow & f(x, y) > \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad \text{(i)}
\end{align*}
\]

If \((x, y) \in B\),
\[
\begin{align*}
& x^2 + y^2 > r^2 \\
\Rightarrow & f(x, y) < \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \quad \text{(ii)}
\end{align*}
\]

From (i) & (ii) we get,
\[
\iint_{x, y \in A} f(x, y) \, dx \, dy > \iint_{x, y \in B} f(x, y) \, dx \, dy
\]

\[
P[x, y \in A] > P[x, y \in B]
\]

This inequality similarly holds for the other quadrants.

\textbullet \quad \text{Problem 7. Show that,}
\[
\frac{1}{\sqrt{2\pi}} \int_{0}^{a} e^{-\frac{z^2}{2}} \, dz < \frac{1}{2} \sqrt{1 - e^{-\frac{2a^2}{\pi}}}
\]

\textbf{Solution:}
\[
P(x, y \in A) = \frac{1}{4} \iint_{x, y \in A} \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \, dx \, dy
\]

\[
P(x, y \in B) = \frac{1}{4} \iint_{x, y \in B} \frac{1}{2\pi \sigma^2} e^{-\frac{r^2}{2\sigma^2}} \, dx \, dy
\]

\[
\Rightarrow \frac{1}{4} \left(1 - e^{-\frac{\pi r^2}{2\sigma^2}}\right) > \frac{1}{2\pi} \left[\int_{x=0}^{a} e^{-x^2/2} \, dx\right]^2
\]

\[
\Rightarrow \frac{1}{\sqrt{2\pi}} \int_{0}^{a} e^{-x^2/2} \, dx < \frac{1}{2} \left(1 - e^{-\frac{\pi r^2}{2\sigma^2}}\right)^{1/2}
\]
MODES OF CONVERGENCE:

A. Convergence in Distribution or in Law:

- **Definition:**
  1. Let \( \{F_n(x)\} \) be a sequence of D.F.'s. If there exists a D.F. \( F(x) \) such that, as \( n \to \infty \), \( F_n(x) \to F(x) \) at every point \( x \) at which \( F(x) \) is continuous, we say that \( \{F_n(x)\} \) converges in distribution or in law to \( F(x) \).

- **Example:** Let \( F_n(x) \) be a sequence of D.F.'s, where
  
  \[
  F_n(x) = \begin{cases} 
  0 & \text{if } x < 0 \\
  \frac{1}{n} & \text{if } 0 \leq x < n \\
  1 & \text{if } x \geq n.
  \end{cases}
  \]

  Does \( \{F_n(x)\} \) converge in distribution?

  **Solution:** Hence, \( \lim_{n \to \infty} F_n(x) = 0 \) if \( x < 0 \)

  \[\lim_{n \to \infty} F_n(x) = 1, \quad 0 \leq x < \infty\]

  Note that \( x \) is a real number, \(-\infty < x < \infty\) and we ignore the point \( x = \infty \).

  Now, \( F(x) \) is a D.F. of a RV degenerated at \( x = 0 \).

  Hence \( \{F_n(x)\} \) converges 'in distribution' or 'weakly' to \( F(x) \).

- **Example:** Let \( F_n(x) \) be a sequence of D.F.'s, where
  
  \[
  F_n(x) = \begin{cases} 
  0 & \text{if } x < -n \\
  \frac{x+n}{2n} & \text{if } -n \leq x < n \\
  1 & \text{if } x \geq n.
  \end{cases}
  \]

  Does \( \{F_n(x)\} \) converge in distribution?

  **Solution:** \( \lim_{n \to \infty} F_n(x) = \frac{1}{2} \), \(-\infty < x < \infty\)

  Clearly, \( F(x) \) is a D.F., as \( F(-\infty) = \frac{1}{2} \neq 0 \) and \( F(\infty) = \frac{1}{2} \neq 1 \).

  Hence, \( \{F_n(x)\} \) does not converge to a D.F.,

  \[\Rightarrow \{F_n(x)\} \not\to F(x)\]


**Convergence in Probability:**

- **Definition:** Let $\{X_n\}$ be a sequence of R.V.'s defined on some probability space $(\Omega, \mathcal{F}, P)$, we say that the sequence $\{X_n\}$ converges in probability to the R.V. $X$, if for every $\epsilon > 0$,
  
  \[ P[|X_n - X| > \epsilon] \to 0 \quad \text{as} \quad n \to \infty . \]

  \[\text{or,} \quad P[|X_n - X| < \epsilon] \to 1 \quad \text{as} \quad n \to \infty \]

  We write, $X_n \xrightarrow{p} X$.

- **Example:**

  Let $\{X_n\}$ be a sequence of R.V.'s each with PMF

  \[ P[X_n = 0] = \frac{1}{2}, \quad P[X_n = 1] = \frac{1}{2} . \]

  Does $\{X_n\}$ converge in probability to some R.V. $X$?

  **Ans:** [Note that, $P[X_n = 0] \to 0$ as $n \to \infty$]

  Now,

  \[ P[|X_n - 0| > \epsilon] = P[X_n > \epsilon] = \begin{cases} 1 - \frac{1}{2}, & 0 < \epsilon < 1 \\ \frac{1}{2}, & \epsilon \geq 1 \end{cases} \]

  \[ = \frac{1}{2} \quad \text{for every} \quad \epsilon > 0, \quad \text{and} \quad n \to \infty \]

  Hence, for every $\epsilon > 0$,

  \[ P[|X_n - 0| > \epsilon] \to 0 \quad \text{as} \quad n \to \infty \]

  \[ \Rightarrow P[|X_n - X| > \epsilon] \to 0 \quad \text{as} \quad n \to \infty \]

  \[ \Rightarrow X_n \xrightarrow{p} X, \quad \text{where} \quad X \text{is a R.V. degenerate at} \quad x = 0 . \]

  Therefore, $X_n \xrightarrow{p} X$. 

Ex. 2. Let $\{x_n\}$ be a sequence of i.i.d. $R(\theta, \Theta)$ r.v.'s. Show that 
\[ X(n) = \max_{i=1}^{n} x_i \]
converges in probability to $\Theta$.

**Solution:**
For any $\epsilon > 0$,
\[
P\left[ \left| x(n) - \Theta \right| < \epsilon \right] = P\left[ \Theta - \epsilon < x(n) < \Theta + \epsilon \right] = F_X(n)(\Theta + \epsilon) - F_X(n)(\Theta - \epsilon)
\]
\[
= \left\{ \begin{array}{ll}
\left( \frac{\theta - \epsilon}{\Theta} \right)^n, & \text{if } 0 < \epsilon < \Theta \\
1 - 0, & \text{if } \epsilon > \Theta
\end{array} \right.
\]
As $0 < \epsilon < \Theta$, $0 < \frac{\Theta - \epsilon}{\Theta} < 1$ and
\[
\lim_{n \rightarrow \infty} \left( \frac{\Theta - \epsilon}{\Theta} \right)^n = 0
\]

**Theorem:**
\[ X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{L} X \]

**Proof:**
Let $F_n(x)$ and $F(x)$ be the cdfs of $X_n$ and $X$, we have,
\[
\{ x_n(x_0) \leq x' \} = \{ x_0 < x', x_n(x_0) \leq x' \} \cup \{ x_0 \geq x', x_n(x_0) \leq x' \}
\]
\[
\leq \{ x_0 \leq x' \} \cup \{ x_0 > x', x_n(x_0) \leq x' \}
\]
Hence, $F_n(x) \leq F(x') + P\left[ x > x', x_n \leq x' \right] = F(x') + \lim_{n \rightarrow \infty} P\left[ x_n > x' \right] \rightarrow 0$ as $n \rightarrow \infty$.

Therefore, $\lim_{n \rightarrow \infty} F_n(x) = F(x')$, $\alpha < \alpha'$

Similarly, by interchanging role $X_n$ and $x$, we get
\[
F(\alpha'') \leq \lim_{n \rightarrow \infty} F_n(\alpha'') \rightarrow x'' < \alpha',
\]
\[
F(\alpha'') \leq \lim_{n \rightarrow \infty} F_n(\alpha') \leq \lim_{n \rightarrow \infty} F_n(x) \leq F(x')
\]

Hence for $\alpha'' < \alpha < \alpha'$,
As \( F(.) \) has only countable number of discontinuity points, we choose \( \alpha \) to be a point of continuity of \( F \), and letting \( \alpha' \downarrow \alpha \) and \( \alpha'' \uparrow \alpha \), we have,

\[
F(\alpha) = \lim_{n \to \infty} F_n(\alpha) \leq \lim_{n \to \infty} F_n(\alpha') = F(\alpha)
\]

\[
\Rightarrow \lim_{n \to \infty} F_n(\alpha) = \lim_{n \to \infty} F(\alpha) = F(\alpha)
\]

\[
\Rightarrow  \lim_{n \to \infty} F_n(\alpha) = F(\alpha), \text{ at all continuity points of } F(x).
\]

Can, \( X_n \xrightarrow{p} x \Rightarrow X_n \xrightarrow{L} x \)

[The convergence in law does not imply the convergence in probability. The converse of the theorem is not true.]

Proof:

Counter-Example: — Let \( \{x_n\} \) be a sequence of identically distributed RV's and let \( \{x, x_n\} \) has the following distribution:

<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1/2</td>
<td>1/2</td>
</tr>
<tr>
<td>1</td>
<td>1/2</td>
<td>0</td>
<td>1/2</td>
</tr>
<tr>
<td>Total</td>
<td>1/2</td>
<td>1/2</td>
<td>1</td>
</tr>
</tbody>
</table>

Clearly, \( x_n \xrightarrow{p} x \), but

\[
p \{ |x_n - x| > 1/2 \} = p \{ |x_n - 1/2| = 1/2 \}
\]

\[
= p \{ x_n = 0, x = 1/2 \} + p \{ x_n = 1/2, x = 0 \}
\]

\[
\xrightarrow{p} \xrightarrow{L} 0, \quad \text{as } \quad \epsilon \to 0
\]

Hence, \( x_n \not\xrightarrow{L} x \). But \( x_n \xrightarrow{p} x \).

\[
\text{Hence, } X_n \not\xrightarrow{L} x, \text{ but } X_n \xrightarrow{p} x.
\]
Ex. 3. Let \( X \sim N(0,1) \). Define \( X_n = \begin{cases} X & \text{if } n = 2m \text{ for some } m \in \mathbb{N} \\ -X & \text{if } n = 2m+1 \end{cases} \) for all \( n \in \mathbb{N} \).

Show that \( X_n \xrightarrow{p} X \), but \( X_n \xrightarrow{d} X \).

**Soln.** As \( X \sim N(0,1) \) and \( X \) is symmetric about '0', \( X \) and \( -X \) have the same distribution.

Hence, \( X_n \sim N(0,1) \) for \( n \in \mathbb{N} \)

and \( X_n \xrightarrow{d} X \),

for any \( \varepsilon > 0 \), when \( n = 2m \),

\[
P \left( |X_n - X| < \varepsilon \right) = P \left( |X| < \frac{\varepsilon}{2} \right) = 2 \Phi \left( \frac{\varepsilon}{2} \right) - 1
\]

\( \xrightarrow{p} 1 \) as \( n \to \infty \).

Hence, \( X_n \xrightarrow{p} X \).

E. \( p \)th mean convergence:

- **Definition:** Let \( \{X_n\} \) be a sequence of R.V.'s such that \( E|X_n|^p < \infty \), for \( p > 0 \), we say that \( \{X_n\} \) converges in the \( p \)th mean to a R.V. \( X \) if \( E|X|^p \to \infty \) and \( E|X_n - X|^p \to 0 \) as \( n \to \infty \).

- **Example:** Let \( \{X_n\} \) be a sequence of R.V.'s such that \( P[X_n = 0] = 1 - \frac{1}{n} \), \( P[X_n = 1] = \frac{1}{n} \), \( m \in \mathbb{N} \). Show that \( \{X_n\} \) converges in 2nd mean to some R.V. \( X \).

**Soln.** \[ E|X_n - 0|^2 = E(X_n^2) = 0 \text{ as } n \to \infty \]

Hence, \( E|X_n - X|^2 \to 0 \) as \( n \to \infty \)

cohere \( X \) is a R.V. degenerate at \( x = 0 \).

\( \xrightarrow{p} \{X_n\} \) converges in 2nd mean to \( X \).
Theorem: \[ \text{If } \lim_{n \to \infty} |x_n - x|^{ \frac{1}{n} } \to 0 \text{ as } n \to \infty \text{ then } x_n \overset{P}{\to} x. \]

Proof: For any \( \varepsilon > 0 \),
\[
P[|x_n - x| > \varepsilon] = P[|x_n - x|^{ \frac{1}{n} } > \varepsilon^{ \frac{1}{n} }] < \frac{E|x_n - x|^{ \frac{1}{n} }}{\varepsilon^{ \frac{1}{n} }}, \quad \text{by Markov's inequality.}
\]

For \( \varepsilon > 0 \),
\[
0 \leq P[|x_n - x| > \varepsilon] < \frac{E|x_n - x|^{ \frac{1}{n} }}{\varepsilon^{ \frac{1}{n} }} \to 0 \text{ as } n \to \infty.
\]
\[
\Rightarrow P[|x_n - x| > \varepsilon] \to 0 \text{ as } n \to \infty,
\]
\[
\Rightarrow x_n \overset{P}{\to} x
\]

Remain: The \( n \)-th mean convergence
\[ \Rightarrow \text{the convergence in probability},
\]
\[ \Rightarrow \text{the convergence in law}.
\]

2) \( \text{The convergence } \Rightarrow \text{in } n\text{-th mean}
\[ \Rightarrow \text{the convergence in probability but the converse is not true}.
\]

Counter Example: Let \( \{x_n\} \) be a sequence of R.V's such that,
\[
P[x_n = 0] = \frac{1}{n}, \quad P[x_n = n] = \frac{1}{n}, \quad n \in \mathbb{N}
\]

For \( \varepsilon > 0 \),
\[
P[|x_n - 0| > \varepsilon] = P[x_n > \varepsilon] = P[x_n = n] = \frac{1}{n} \to 0 \text{ as } n \to \infty.
\]

Hence, \( x_n \overset{P}{\to} 0 \)
\[
\Rightarrow x_n \overset{P}{\to} x, \text{ where } x \text{ is a R.V degenerate at } x = 0.
\]

But,
\[
E|x_n - x| = E|x_n - 0|^{\frac{1}{n}} = E(x_n^{\frac{1}{n}})
\]
\[
= n^{\frac{1}{n}} (1 - \frac{1}{n}) + n^{\frac{1}{n}} \frac{1}{n} = n \to \infty \text{ as } n \to \infty
\]
\[
i.e., \to 0.
\]

Hence, \( x_n \overset{P}{\to} x \) but \( \{x_n\} \) does not converge in 2nd mean to \( x \).
WEAK LAW OF LARGE NUMBERS (WLLN):

Let \( \{X_n\} \) be a sequence of R.V.s, let
\[
S_n = \sum_{k=1}^{n} X_k, \quad n \in \mathbb{N},
\]
we say that \( \{X_n\} \) obeys the weak law of large numbers (WLLN) with respect to the sequence \( \{b_n\} \)
\( b_n \to 0 \) and \( b_n \to \infty \), if \( \forall \varepsilon > 0 \) there exists an infinite sequence \( \{a_n\} \)
of real numbers such that
\[
\frac{S_n - a_n}{b_n} \xrightarrow{p} 0.
\]
Here, \( a_n \) is called the centering constant and \( b_n \) is called the normalizing constant.

Chebyshev's WLLN:

Let \( \{X_n\} \) be a sequence of independent R.V.s such that
\[
E(X_n) = \mu_n \quad \text{and} \quad \text{Var}(X_n) = \sigma_n^2 < \infty,
\]
then
\[
\lim_{n \to \infty} \left( \frac{\sum_{k=1}^{n} X_k}{n} \right) = \mu,
\]
\[
\Rightarrow \frac{X_n - \mu_n}{\sigma_n} \xrightarrow{p} 0
\]
where \( X_n = \frac{1}{n} \sum_{k=1}^{n} X_k \), \( \mu_n = \frac{1}{n} \sum_{k=1}^{n} \mu_k \).

So \( \{X_n\} \) obeys WLLN.

Proof:

\[
E(X_n) = \frac{1}{n} \sum_{k=1}^{n} E(X_k) = \frac{1}{n} \sum_{k=1}^{n} \mu_k = \mu_n,
\]
\[
\text{Var}(X_n) = \text{Var} \left( \frac{1}{n} \sum_{k=1}^{n} X_k \right) = \frac{1}{n^2} \sum_{k=1}^{n} \text{Var}(X_k) = \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2,
\]
as \( X_k \)'s are independent.

For every \( \varepsilon > 0 \),
\[
0 \leq P \left[ \left| \frac{X_n - \mu_n}{\sigma_n} \right| > \varepsilon \right] < \frac{\text{Var}(X_n)}{\varepsilon^2} = \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 \to 0
\]
provided \( \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 \to 0 \) as \( n \to \infty \).

Hence,
\[
\lim_{n \to \infty} \frac{\sum_{k=1}^{n} \sigma_k^2}{n^2} = 0.
\]

\[
\Rightarrow P \left[ \left| \frac{X_n - \mu_n}{\sigma_n} \right| > \varepsilon \right] \to 0 \quad \text{as} \quad n \to \infty \quad \forall \varepsilon > 0
\]
\[
\Rightarrow \frac{X_n - \mu_n}{\sigma_n} \xrightarrow{p} 0
\]
\[
\Rightarrow \frac{S_n - \sum_{k=1}^{n} \mu_k}{n} \xrightarrow{p} 0
\]
\[
\Rightarrow \frac{S_n - a_n}{b_n} \xrightarrow{p} 0
\]
where \( a_n = \sum_{k=1}^{n} \mu_k \), \( b_n = n \to 0 \) and \( b_n \to \infty \).

\( \Rightarrow \{X_n\} \) obeys WLLN.
Example 1. Examine whether the WLLN holds for the following sequences $\{X_n\}$ of independent R.V.'s:

$$P[X_n = -2^n] = 2^{-2n-1} = P[X_n = 2^n]$$
$$P[X_n = 0] = 1 - 2^{-2n}$$

$$P[X_n = -n] = \frac{1}{2} = P[X_n = n]$$

**Soln.**

$$\mu_k = E(X_k) = (-2^k) \cdot 2^{-2k-1} + (2^k) \cdot 2^{-2k-1} + 0 \cdot (1 - 2^{-2k}) = 0$$

and

$$Var(X_k) = \sigma_k^2 = E(X_k^2)$$

$$= (-2^k)^2 \cdot 2^{-2k-1} + (2^k)^2 \cdot 2^{-2k-1} + 0$$

$$= 1, \quad k \in \mathbb{N}$$

Now,

$$\frac{1}{n} \sum_{k=1}^{n} \sigma_k^2 = \frac{1}{n} \sum_{k=1}^{n} 1 = \frac{1}{n} \to 0 \quad \text{as} \quad n \to \infty.$$

Hence, $\{X_n\}$ obeys WLLN, by Chebyshev's WLLN.

ii)

Hence, $\mu_k = 0$ and

$$\sigma_k^2 = \frac{1}{k}, \quad k \in \mathbb{N}$$

Now,

$$\frac{1}{n} \sum_{k=1}^{n} \sigma_k^2 = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} < \frac{c}{n}$$

is a convergent $p$-series,

$$\Rightarrow \frac{1}{n} \sum_{k=1}^{n} \frac{1}{k} = \frac{\ln n}{n} < \frac{c}{n}, \quad \text{a finite quantity}$$

Hence, $\{X_n\}$ obeys WLLN, by Chebyshev's WLLN.

Ex. 2. Let $P[X_n = -n^P] = \frac{1}{2} = P[X_n = n^P]$. Show that WLLN holds for the sequence $\{X_n\}$ of independent R.V.'s if $P < \frac{1}{2}$.

**Soln.**

Hence,

$$\mu_k = E(X_k) = 0$$

$$\sigma_k^2 = V(X_k) = E(X_k^2) = (-k^P)^2 \cdot \frac{1}{2} + (k^P)^2 \cdot \frac{1}{2} = k^{2P}$$

$$k \in \mathbb{N}$$

Now,

$$\frac{1}{n} \sum_{k=1}^{n} \sigma_k^2 = \frac{1}{n} \sum_{k=1}^{n} k^{2P} < \frac{1}{n} \int_{0}^{\infty} x^{2P} \, dx$$

$$= \frac{n^{2P+1} - 1}{n(2P+1)}$$
Now,
\[
0 \leq \frac{1}{n^p} \sum_{k=1}^{n} \sigma_k^p \leq \frac{n^{2p+1}-1}{n^p(2p+1)} \leq \frac{n^{2p}}{2p+1} \rightarrow 0 \text{ as } n \rightarrow \infty.
\]

\[\Rightarrow \text{ if } p < \frac{1}{2}, \frac{1}{n^p} \sum_{k=1}^{n} \sigma_k^p \rightarrow 0 \text{ as } n \rightarrow \infty.\]

Hence, \( \{X_n\} \) obeys \( \text{WLLN} \) if \( p < \frac{1}{2} \).

**Ex. 3.** Decide whether \( \text{WLLN} \) holds for the sequence \( \{X_n\} \) of independent R.V.'s.

i) \( P[X_n = \pm 2^n] = \frac{1}{2^n} \)

\[\Rightarrow P[X_n = -2^n] = \frac{1}{2} = P[X_n = 2^n].\]

ii) \( P[X_n = n] = P[X_n = -n] = \frac{1}{2n} \)

\[P[X_n = 0] = 1 - \frac{1}{2n}.\]

**SOLN.**

1) \( \mu_k = 0, \)
and \( \sigma_k^p = E(X_k) = E(X_k^p) = 2 - 2^k, k \in \mathbb{N} \)

Now,
\[
\frac{1}{n^p} \sum_{k=1}^{n} \sigma_k^p = \frac{1}{n^p} \sum_{k=1}^{n} 2 - 2^k = \frac{1}{n^p} \sum_{k=1}^{n} \left( \frac{1}{2} \right)^k \frac{1}{1 - \frac{1}{4}} = \frac{1}{3} \cdot \frac{1}{n^p} \left( 1 - \left( \frac{1}{4} \right)^n \right) < \frac{1}{3n^p}.
\]

Hence,
\[
\lim_{n \rightarrow \infty} \frac{1}{n^p} \sum_{k=1}^{n} \sigma_k^p = 0
\]

\[\Rightarrow \{X_n\} \text{ obeys WLLN by Chebyshev's WLLN.}\]

\[\mu_k = 0, \]
and \( \sigma_k^p = E(X_k) = E(X_k^p) = 2 - 2^k, k \in \mathbb{N}.\)

Now,
\[
\frac{1}{n} \sum_{k=1}^{n} \sigma_k^p = \frac{1}{n} \sum_{k=1}^{n} 2 - 2^k = \frac{1}{n} \sum_{k=1}^{n} \left( \frac{1}{2} \right)^k \frac{1}{1 - \frac{1}{4}} = \frac{1}{3} \cdot \frac{1}{n} \left( 1 - \left( \frac{1}{4} \right)^n \right) = \frac{1}{3n} \left( 1 - \left( \frac{1}{4} \right)^n \right) < \frac{1}{3n}.
\]

Hence, \( \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^{n} \sigma_k^p = 0 \)

\[\Rightarrow \sum_{k=1}^{n} \sigma_k^p \rightarrow 0 \text{ as } n \rightarrow \infty.
\]

For large \( n, \)
\[
\frac{1}{n} \sum_{k=1}^{n} \left( \frac{k}{n} \right)^{3/2} \approx \frac{2n^{5/2}}{5}
\]

\[\Rightarrow \sum_{k=1}^{n} k^{3/2} \approx \frac{2n^{5/2}}{5}
\]

For large \( n, \)
\[
\frac{1}{n} \sum_{k=1}^{n} k^{3/2} = \frac{1}{n} \cdot \frac{2n^{5/2}}{5} = \frac{2}{5} n^{5/2} \rightarrow 0
\]

As \( \frac{1}{n} \sum_{k=1}^{n} \sigma_k^p \rightarrow 0 \text{ as } n \rightarrow \infty, \)

we cannot draw any conclusion by Chebyshev's WLLN, whether WLLN holds or not.
Markov's WLLN: Let \( \{X_n\} \) be a sequence of RV's such that \( E(X_n) = \mu_n \) and \( \gamma(X_n) = \sigma_n < \infty \), \( n \in \mathbb{N} \), then

\[
X_n - \bar{X}_n \xrightarrow{p} 0
\]

provided \( \frac{1}{n\nu} \text{Var} \left( \sum_{k=1}^{n} X_k \right) \to 0 \) as \( n \to \infty \).

i.e. \( \{X_n\} \) obeys WLLN, provided \( \frac{1}{n\nu} \text{Var} \left( \sum_{k=1}^{n} X_k \right) \to 0 \) as \( n \to \infty \).

**Proof:** For any \( \epsilon > 0 \),

\[
0 \leq P \left[ |X_n - \bar{X}_n| > \epsilon \right] \leq \frac{1}{\epsilon^2} \text{Var}(X_n)
\]

provided \( \frac{1}{n\nu} \text{Var} \left( \sum_{k=1}^{n} X_k \right) \to 0 \) as \( n \to \infty \).

**Remark:** Chebyshev's WLLN is a particular case of Markov's WLLN.

**Ex. 4.1** Let \( \{X_n\} \) be a sequence of RV's with common finite variances. Suppose \( \rho_{X_i, X_j} \leq \epsilon \) for \( i \neq j \). Prove that WLLN holds for \( \{X_n\} \).

**Solution:** Note that,

\[
\frac{1}{n\nu} \text{Var} \left( \sum_{k=1}^{n} X_k \right) = \frac{1}{n\nu} \left( \sum_{k=1}^{n} \gamma(X_k) + \sum_{k \neq \ell} \rho_{X_k, X_\ell} \sigma_k \sigma_\ell \right)
\]

Hence,

\[
0 \leq \frac{1}{n\nu} \gamma \left( \sum_{k=1}^{n} X_k \right) \leq \frac{1}{n\nu} \sum_{k=1}^{n} \gamma(X_k) \to 0 \quad \text{as} \quad n \to \infty
\]

Hence, for every \( \epsilon > 0 \),

\[
0 \leq P \left[ |X_n - \bar{X}_n| > \epsilon \right] \leq \frac{\gamma \left( \sum_{k=1}^{n} X_k \right)}{\epsilon^2} \to 0 \quad \text{as} \quad n \to \infty
\]

\[
P \left[ |X_n - \bar{X}_n| > \epsilon \right] \to 0 \quad \text{as} \quad n \to \infty
\]

\[
\Rightarrow X_n - \bar{X}_n \xrightarrow{p} 0
\]

\[
\Rightarrow \{X_n\} \text{ obeys WLLN.}
\]
Bernoulli's Law of Large Numbers: Let $\frac{1}{n}$ be the number of occurrences of an event $A$ in $n$ independent trials, and $p$ be the probability of occurrence of an event $A$ in each trial. Then for every $e > 0$; 

$$P\left[\left|\frac{\hat{p}}{n} - p\right| < e\right] \rightarrow 1, \text{ as } n \rightarrow \infty.$$ 

In other words, the sequence of relative frequency of the event $A$, $\frac{\hat{p}}{n}$ converges in probability to $p$.

**Proof:** Consider the occurrence of the event $A$ as a success. Then $\frac{1}{n}$, the number of successes in $n$ independent Bernoulli trials, follows Bin$(n, p)$.

Hence, $E(\hat{p}) = np$, $V(\hat{p}) = np(1-p)$.

$$E\left(\frac{\hat{p}}{n}\right) = p, \quad V\left(\frac{\hat{p}}{n}\right) = \frac{p(1-p)}{n}.$$ 

For every $e > 0$,

$$P\left[\left|\frac{\hat{p}}{n} - p\right| > e\right] \leq \frac{E\left(\left(\frac{\hat{p}}{n} - p\right)^2\right)}{e^2} = \frac{V\left(\frac{\hat{p}}{n}\right)}{e^2} = \frac{p(1-p)}{ne^2}.$$ 

Hence, for every $e > 0$,

$$0 \leq P\left[\left|\frac{\hat{p}}{n} - p\right| > e\right] \leq \frac{p(1-p)}{ne^2} \leq \frac{1}{4ne^2} \rightarrow 0 \text{ as } n \rightarrow \infty.$$ 

$$\therefore P\left[\left|\frac{\hat{p}}{n} - p\right| < e\right] \rightarrow 1 \text{ as } n \rightarrow \infty.$$ 

Hence, by AM $\geq$ GM,

$$\frac{P + (1-P)}{2} \geq \sqrt{P(1-P)}.$$ 

$$\Rightarrow \frac{1}{2} \geq \sqrt{P(1-P)}.$$ 

$$\Rightarrow \frac{1}{4} \geq P(1-P).$$ 

**Remark:** We have $P\left[\left|\frac{\hat{p}}{n} - p\right| < e\right] \rightarrow 1$, as $n \rightarrow \infty$, for every $e > 0$,

$\Rightarrow$ for large $n$, the values of $\frac{\hat{p}}{n}$ are very close to $p$ with probability $\approx 1$.

$\Rightarrow$ for large $n$, $p \leq \frac{\hat{p}}{n}$, which is nothing but Statistical definition of probability.

Therefore, the Bernoulli's Law of Large Numbers is the foundation of statistical definition of probability.
Theorem: A necessary & sufficient condition for WLLN.

Let \( \{X_n\} \) be a sequence of RV's, Define,

\[ Y_n = \frac{1}{n} \sum_{k=1}^{n} X_k \]

A necessary & sufficient condition for the sequence \( \{X_n\} \) to satisfy the WLLN is that

\[ E\left( \frac{Y_n}{1+Y_n} \right) \to 0 \text{ as } n \to \infty. \]

Proof: Note that,

\[ \frac{y}{1+y} > \frac{e^y}{1+e^y} \]

\[ \Rightarrow y > e^y \Rightarrow 1/y < e^y \]

For \( \varepsilon > 0 \),

\[ P\left[ |Y_n| > \varepsilon \right] = P\left[ \frac{Y_n}{1+Y_n} > \frac{e^y}{1+e^y} \right] \]

\[ \leq \frac{e^y}{1+e^y} \text{ by Markov's inequality.} \]

If \( E\left( \frac{Y_n}{1+Y_n} \right) \to 0 \text{ as } n \to \infty, \)

\[ P\left[ |Y_n| > \varepsilon \right] \to 0 \text{ as } n \to \infty \]

\[ \Rightarrow Y_n \xrightarrow{p} 0 \text{ i.e. } X_n \xrightarrow{p} 0 \]

i.e. \( \{X_n\} \) obeys WLLN.

Now,

\[ E\left( \frac{Y_n}{1+Y_n} \right) = \int_{-\infty}^{\infty} \frac{y}{1+y} \cdot dF_n(y) \]

\[ \leq \int_{|y| \leq \varepsilon} \frac{y}{1+y} \cdot dF_n(y) + \int_{|y| > \varepsilon} \frac{y}{1+y} \cdot dF_n(y) \]

\[ \leq \varepsilon e^\varepsilon \cdot \int_{|y| \leq \varepsilon} dF_n(y) + \varepsilon \cdot \int_{|y| > \varepsilon} dF_n(y) \]

\[ \leq \varepsilon e^\varepsilon \cdot + \varepsilon \cdot \int_{|y| > \varepsilon} dF_n(y) \]

Hence,

\[ E\left( \frac{Y_n}{1+Y_n} \right) \leq e^\varepsilon \cdot 1 + P\left[ |Y_n| > \varepsilon \right] \]

If \( \{X_n\} \) obeys WLLN, then for every \( \varepsilon > 0, \)

\[ P\left[ |Y_n| > \varepsilon \right] \to 0 \text{ as } n \to \infty \]

\[ \Rightarrow E\left( \frac{Y_n}{1+Y_n} \right) \leq e^\varepsilon \text{ for large } n. \]

\[ \Rightarrow E\left( \frac{Y_n}{1+Y_n} \right) \to 0 \text{ as } n \to \infty. \]
**Central Limit Theorem:**

For a sequence \( \{X_n\} \) of i.i.d. R.V.'s, we have \( X_n \xrightarrow{d} \mu \), provided \( \mu \) exists. Hence, WLLN holds for i.i.d. sequence \( \{X_n\} \) of R.V.'s, provided mean \( \mathbb{E}(X) = \mu \) exists. But this gives no idea as to how the distribution of \( \bar{X}_n \) can be approximated in large samples. Hence, we consider the condition under which the distribution of \( S_n = \sum_{k=1}^{n} X_k \) or \( \bar{X}_n \) converges to normal distribution.

**Definition:**

If the distribution of a R.V. \( Y_n \) depends on a parameter \( n \), and if there exists two quantities \( a_n \) and \( b_n \) (which may or may not depend on \( n \)) such that

\[
\lim_{n \to \infty} \frac{Y_n - a_n}{b_n} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} dt, \text{ for all } y \in \mathbb{R}
\]

then we say that \( Y_n \) is asymptotically normally distributed with mean \( a_n \) and variance \( b_n \). We also say that \( \frac{Y_n - a_n}{b_n} \) follows the central limit theorem on normal convergence.

**Notation:**

\[
Y_n \xrightarrow{d} X \sim \mathcal{N}(0,1).
\]

\[
\Rightarrow Y_n \sim \mathcal{N}(a_n/b_n).
\]

**Lindeberg-Levy CLT [i.i.d. case]:**

Let \( \{X_n\} \) be a sequence of i.i.d. R.V.'s with common mean \( \mu \) and finite variance \( \sigma^2 \). Let, \( S_n = \sum_{k=1}^{n} X_k, n \in \mathbb{N} \)

Then, for any \( x \in \mathbb{R} \),

\[
p\left[ \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \leq x \right] \xrightarrow{d} \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt, \text{ as } n \to \infty.
\]

\[
\Rightarrow p\left[ \frac{X_n - \mu}{\sigma\sqrt{n}} \leq x \right] \xrightarrow{d} \Phi(x), \text{ as } n \to \infty
\]

**Remark:**

For a sequence \( \{X_n\} \) of i.i.d. R.V.'s, by Lindeberg-Levy CLT, we have,

\[
\frac{S_n - \mathbb{E}(S_n)}{\sqrt{\mathbb{V}(S_n)}} \xrightarrow{d} X \sim \mathcal{N}(0,1)
\]

\[
\Rightarrow \frac{X_n - \mathbb{E}(X_n)}{\sqrt{\mathbb{V}(X_n)}} \xrightarrow{d} X \sim \mathcal{N}(0,1)
\]

Hence, \( Y_n = \frac{S_n}{\sqrt{\mathbb{V}(X_n)}} \), \( a_n = \mathbb{E}(Y_n) \), \( b_n = \sqrt{\mathbb{V}(Y_n)} \).
Ex. 1. Let $Y_n$ be the number of independent trials necessary to get $r$th success, where $p$ is the probability of success in each trial. Find \( \lim_{n \to \infty} p \left[ \frac{B_n - n}{\sqrt{n}} \leq y \right] \).

**Soln.** Let \( \{X_n\} \) be a sequence of i.i.d. R.V.'s where each \( X_n \) denotes the number of trials required to get the 1st success in a sequence of independent Bernoulli trials. Clearly, \( X_n \sim \text{Geo}(p) \), \( n \in \mathbb{N} \).

Then, \( Y_n = \sum_{k=1}^{r} X_k \) = Number of trials required to get the \( r \)th success.

Note that, \( E(X_n) = \frac{1}{p} \), \( V(X_n) = \frac{1}{p^2} \), a finite quantity. \( \therefore \)

Applying Lindeberg - Levy's form, the sequence \( \{X_n\} \) of i.i.d. R.V.'s is said to be of finite variance, \( \frac{Y_n - E(Y_n)}{\sqrt{V(Y_n)}} \rightarrow \Phi(y) \), as \( n \to \infty \), \( y \in \mathbb{R} \).

\[ \Rightarrow \quad \frac{Y_n - \frac{n}{p}}{\sqrt{\frac{n}{p^2}}} \rightarrow \Phi(y) \], as \( n \to \infty \).

\[ \Rightarrow \quad \lim_{n \to \infty} p \left[ \frac{B_n - n}{\sqrt{n}} \leq y \right] = \Phi(y) \]

Ex. 2. Using CLT, evaluate the limit \( \lim_{n \to \infty} \sum_{k=0}^{n} \frac{(-1)^{n-k} \cdot n^k}{k!} \).

**Soln.** Let \( \{X_n\} \) be a sequence of i.i.d. Poisson variable with mean \( \lambda = 1 \).

Note that, \( E(X_n) = \lambda(X_n) = 1 \), finite \( n \in \mathbb{N} \).

Also, \( S_n = \sum_{k=1}^{n} X_k \sim \text{P}(\lambda) \).

By Lindeberg - Levy's CLT,
\[ \lim_{n \to \infty} p \left[ \frac{S_n - E(S_n)}{\sqrt{V(S_n)}} \leq x \right] = \Phi(x) \quad \forall \alpha \in \mathbb{R} \]

\[ \Rightarrow \quad \lim_{n \to \infty} p \left[ \frac{S_n - n}{\sqrt{n}} \leq x \right] = \Phi(x) \quad \forall \alpha \in \mathbb{R} \]

For \( x = 0 \), \( \lim_{n \to \infty} p \left[ S_n \leq n \right] = \Phi(0) \).

\[ \Rightarrow \quad \lim_{n \to \infty} \sum_{k=0}^{n} p \left[ S_n = k \right] = \frac{1}{2} \]

\[ \Rightarrow \quad \lim_{n \to \infty} \sum_{k=0}^{n} \frac{(-1)^{n-k} \cdot n^k}{k!} = \frac{1}{2} \]
Ex. 3. Using CLT, evaluate \( \lim_{n \to \infty} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \, dx \).

**Solution:** Let \( \{X_n\} \) be a sequence of i.i.d. exponential R.V.'s with mean \( 1 \).

Then, \( S_n = \sum_{k=1}^{n} X_k \sim \Gamma(n) \).

\[ S_n = \frac{1}{\sqrt{2\pi n}} e^{-\frac{s^2}{2n}} \]

Ex. 4. Let \( \{X_n\} \) be a sequence of i.i.d. R.V.'s with Gamma distribution.

\[ \text{Ex. 4. Evaluate: } \lim_{n \to \infty} \frac{1}{2n} \Gamma\left(\frac{n}{2}\right) \int_{0}^{\infty} e^{-t/2} \, dt. \]

**Solution:** Let \( \{X_n\} \) be a sequence of i.i.d. R.V.'s following Gamma \( \left( \frac{1}{2}, \frac{1}{2} \right) \).

Then, \( S_n = \sum_{k=1}^{n} X_k \sim \Gamma\left( \frac{n}{2}, \frac{1}{2} \right) \).

\[ E(S_n) = n, \quad \gamma(S_n) = 2n. \]

By Lindeberg-Lévy CLT,

\[ \lim_{n \to \infty} \frac{S_n - E(S_n)}{\sqrt{\gamma(S_n)}} = 0. \]
De-Moivre and Laplace Limit Theorem:

If \( \{X_n\} \) be a sequence of i.i.d. Bernoulli RV's with probability \( p \) of success in each trial, then:

\[
\frac{S_n - np}{\sqrt{np(1-p)}} \overset{d}{\to} N(0,1) \quad \text{as} \quad n \to \infty, \quad \text{where} \quad S_n = \sum_{k=1}^{n} X_k.
\]

Proof:

Let \( Z_n = \frac{S_n - np}{\sqrt{np(1-p)}} \).

The MGF of \( Z_n \) is

\[
M_{Z_n}(t) = \mathbb{E}[e^{tZ_n}] = e^{t \left( \frac{S_n - np}{\sqrt{np(1-p)}} \right)}
\]

\[
= e^{t \left( \frac{S_n - np}{\sqrt{np(1-p)}} \right)} \mathbb{E}[e^{\frac{t}{\sqrt{np(1-p)}} S_n}]
\]

\[
= e^{\frac{tnp}{\sqrt{np(1-p)}}} \mathbb{E}[\left( e^{\frac{t}{\sqrt{np(1-p)}} S_n} \right)^n]
\]

\[
= \mathbb{E}\left[ e^{\frac{tnp}{\sqrt{np(1-p)}} (n-3/2)} \left( 1 + \frac{t}{\sqrt{np(1-p)}} S_n \right)^n \right]
\]

\[
= \left[ 1 + \frac{t}{\sqrt{np(1-p)}} (n-3/2) \right]^n
\]

Now, \( \ln M_{Z_n}(t) = n \ln e \left( 1 + \frac{t}{\sqrt{np}} (n-3/2) \right) \)

\[
= n \left( \frac{t}{\sqrt{np}} + o\left( n^{-3/2} \right) \right)
\]

\[
= \frac{t}{2} + o\left( n^{-1/2} \right) \rightarrow \frac{t}{2} \quad \text{as} \quad n \to \infty.
\]

Hence, \( \lim_{n \to \infty} M_{Z_n}(t) = e^{t^2/2} = M(t) \), which is the MGF of \( N(0,1) \).

By uniqueness of MGF,

\[
Z_n = \frac{S_n - np}{\sqrt{np(1-p)}} \overset{d}{\to} N(0,1) \quad \text{as} \quad n \to \infty.
\]
Normal Approximation to Poisson:

[CLT for a sequence of i.i.d. Poisson RV's]

If \( \{X_n\} \) be a sequence of i.i.d. RV's each following \( \text{P}(\lambda) \) distribution, then,

\[
\frac{S_n - n\lambda}{\sqrt{n\lambda}} \xrightarrow{L} Z \sim N(0,1)
\]

**Proof:**

Let, \( Z_n = \frac{S_n - n\lambda}{\sqrt{n\lambda}} \).

\[
M_{Z_n}(t) = E\left[e^{t Z_n}\right] = e^{-t\sqrt{n\lambda}} E\left[e^{\frac{t}{\sqrt{n\lambda}} S_n}\right] = e^{-t\sqrt{n\lambda}} \left[M_{S_n}\left(\frac{t}{\sqrt{n\lambda}}\right)\right] = e^{-t\sqrt{n\lambda}} n\lambda \left(e^{rac{t}{\sqrt{n\lambda}}} - 1\right) \quad \text{as} \quad S_n \sim \text{P}(n\lambda)
\]

\[
e^{-t\sqrt{n\lambda} + n\lambda \left(\frac{t}{\sqrt{n\lambda}} + \frac{1}{2} \frac{t^2}{n\lambda} + o(n^{-1/2})\right)} = e^{\frac{t^2}{2} + o(n^{-1/2})} \xrightarrow{n \to \infty} e^{\frac{t^2}{2}} \quad \text{as} \quad n \to \infty
\]

Hence, \( \lim_{n \to \infty} M_{Z_n}(t) = e^{\frac{t^2}{2}} = M(t) \), which is the mgf of \( N(0,1) \) distribution.

By uniqueness of mgf, \( Z_n \xrightarrow{L} Z \sim N(0,1) \).

\[\square\]

[CLT for a sequence of i.i.d. Gamma (\( \theta, 1 \)), i.e., \( \text{Exp}(\theta) \) RV's]

If \( \{X_n\} \) be a sequence of i.i.d. Exponential RV's each mean \( \theta \), then

\[
\frac{S_n - n\theta}{\sqrt{n\theta^2}} \xrightarrow{L} Z \sim N(0,1) \quad \text{as} \quad n \to \infty
\]

**Proof:**

Let, \( Z_n = \frac{S_n - n\theta}{\sqrt{n\theta^2}} \).

\[
M_{Z_n}(t) = E\left[e^{t Z_n}\right] = e^{-t\theta}\left[e^{\frac{t}{\theta}} S_n\right] = e^{-t\theta} M_{S_n}\left(\frac{t}{\theta}\right) = e^{-t\theta} \left(1 - \frac{t}{\theta}\right)^{-n}
\]

\[\text{[Hence, } S_n = \sum_{k=1}^{n} X_k \sim \text{Gamma}(\theta, n)\text{; } M_{S_n}(t) = (1 - \theta t)^{-n}, \text{if } t < \frac{1}{\theta}\text{]}\]
Now, \( \ln M_{Z_n}(t) = -t \ln n - n \ln \left(1 - \frac{t}{\sqrt{n}}\right) \)
\[= -t \ln n - n \left\{ \frac{t}{\sqrt{n}} + \frac{t^2}{2n} + o\left(n^{-3/2}\right) \right\} \]
\[= -t \ln n + t \ln n + \frac{t^2}{2} + o\left(n^{-3/2}\right) \rightarrow t^2/2 \quad \text{as} \ n \rightarrow \infty. \]

Hence, \( \lim_{n \rightarrow \infty} M_{Z_n}(t) = e^{t^2/2} = M_{Z^2}(t) \), which is the MGF of \( Z \sim N(0,1) \).

By uniqueness of MGF; \( Z_n \xrightarrow{d} \frac{S_m - n\mu}{\sqrt{mn}} \rightarrow Z \sim N(0,1) \)

**Relationship between CLT and WLLN:**

CLT is a generalization of WLLN for a sequence of i.i.d. R.V.'s with finite variance.

Let \( \{X_n\} \) be a sequence of i.i.d. R.V.'s with common mean \( \mu \) and variance \( \sigma^2 < \infty \).

By Lindeberg-Lévy CLT,

\[ \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} \xrightarrow{d} Z \sim N(0,1) \]

Now, \( P\left[|\overline{X}_n - \mu| < \epsilon\right] = P\left[\left|\frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}\right| < \frac{\epsilon \sqrt{n}}{\sigma}\right] \)
\[= P\left[|Z| < \frac{\epsilon \sqrt{n}}{\sigma}\right], \text{ for large } n. \]
\[= 2\Phi\left(\frac{\epsilon \sqrt{n}}{\sigma}\right) - 1 \rightarrow 2\Phi(+\infty) - 1 = 1. \]

as \( n \rightarrow \infty \), for every \( \epsilon > 0 \).

Hence, \( \overline{X}_n \xrightarrow{p} \mu \).

\( \iff \{X_n\} \) obeys WLLN.

Hence, CLT is stronger than WLLN for a sequence of i.i.d. R.V.'s with finite variance.

But for the sequence \( \{X_n\} \) of independent R.V.'s, CLT may hold but the WLLN may not hold.
Further, WLLN and CLT:

A. KINCHIN'S WLLN: \( \Rightarrow \) If \( \{X_n\} \) is a sequence i.i.d. R.V.'s, then \( \mu = E(X_1) \) exists.

\[ \Rightarrow \frac{X_n}{n} \xrightarrow{P} \mu, \text{i.e., } \{X_n\} \text{ obeys WLLN.} \]

Ex.1) Check whether WLLN holds for the following sequence of i.i.d. R.V.'s (a) cdf PDF, \( f(x) = \begin{cases} \frac{1+x}{1^2+x^2}, & x > 1 \\ 0, & \text{otherwise} \end{cases} \)

where \( \delta > 0 \).

(b) with PMF \( P[X = (-1)^k, k] = \frac{1}{k^2} \), \( k \geq 1 \)

(c) with PMF \( P[X = 2^{-k} \log k] = 2^{-k} \), \( k \geq 1 \).

Soln. \( \Rightarrow \)

(a) \( E|X_1| = \int_1^{\infty} |x| \cdot \frac{1+x}{x^2+x^2} \, dx \)

\[ = (1+x) \int_1^{\infty} \frac{1}{x+1} \, dx \]

\[ = (1+x) \lim_{a \to \infty} \int_1^{a} \frac{1}{x+1} \, dx \]

\[ = (1+x) \lim_{a \to \infty} \left[ \ln \frac{a}{a+1} \right] \]

\[ = (1+x) \lim_{a \to \infty} \left[ 1 - \frac{1}{a} \right] \]

\[ = 1 + \frac{x}{x}, \text{ for } x > 0 \]

Hence \( E(X) \) exists and \( E(X_1) = \frac{1+x}{x} \).

By KINCHIN'S WLLN, \( \frac{X_n}{n} \xrightarrow{P} \frac{1+x}{x} = E(X_1) \)

i.e., \( \{X_n\} \) obeys WLLN.
(c) \[ E|X_1| = E(X_1) \text{ as } X_1 \text{ is non-negative, } \Rightarrow \log 2 = 2 \log k \]
\[ E|X_1| = \frac{1}{q_{\log k}} \]
\[ E|X_1| = \frac{1}{k_{\log k}} \text{, which is } p \text{-series with } p = \log_4 e > 1 \]

Hence, \( E(X_1) = M \) exists.

By Kinetin's WLLN, \( \bar{X}_n \to E(X) \),

i.e. \( \{X_n\} \) obeys WLLN.

3. Liapounov's CLT:

Let \( \{X_n\} \) be a sequence of independent R.V.'s

with \( E(X_1) = \mu_1 \), \( V(X_1) = \sigma_1^2 \) and,

\[ b_n = E|X_1 - \mu_1|^3 < \infty, \forall i \]

Define, \( B_n = \sum_{i=1}^{n} \frac{X_i}{b_i} \) and \( \sigma_i = \sqrt{\frac{\sigma_i^2}{b_i}} \),

If \( \lim_{n \to \infty} \frac{B_n}{\sigma_i} = 0 \), then

\[ \frac{S_n - \mu \frac{\sigma_i}{b_i}}{\sqrt{\frac{\sigma_i^2}{b_i}}} \sim N(0,1) \]

[\[ Let \{X_n\} be a sequence of independent R.V.'s
with \( E(X_1) = \mu_1 \), \( V(X_1) = \sigma_1^2 \), and

\[ E|X_1 - \mu_1|^3 < \infty \]

for some \( \delta > 0, \forall i \), Define, \( b_n = \sum_{i=1}^{n} \frac{X_i}{b_i} \),

and \( \sigma_i = \sqrt{\frac{\sigma_i^2}{b_i}} \).

If \( \lim_{n \to \infty} \frac{B_n}{\sigma_i} = 0 \), then

\[ \frac{S_n - \mu \frac{\sigma_i}{b_i}}{\sqrt{\frac{\sigma_i^2}{b_i}}} \sim N(0,1) \]
Ex. 1. Examine if the CLT and the WLLN holds for the sequence of independent R.V.'s $X_n$, where

\[ P[X_n = -n] = P[X_n = n] = \frac{1}{2n^2}, \]

\[ P[X_n = 0] = 1 - \frac{1}{n}. \]

\[ \mu = E(X_k) = 0 \]

\[ \sigma_k = k^{3/2} \]

\[ \psi_k = E|X_k - \mu|^3 = k^{5/2}. \]

Then,

\[ \sigma = \sum_{k=1}^{n} k^{3/2} = n^{5/2} \]

\[ \psi = \sum_{k=1}^{n} \psi_k = n^{13/2}. \]

\[ \lim_{n \to \infty} \left( \frac{\sum_{k=1}^{n} \psi_k}{n^{3/2}} \right) = \lim_{n \to \infty} \left( \frac{\sum_{k=1}^{n} k^{5/2}}{n^{13/2}} \right) \]

\[ = \lim_{n \to \infty} \left( \frac{\int_0^n x^{3/2} \, dx}{n^{3/2}} \right) \]

\[ = \lim_{n \to \infty} \frac{2n^{5/2}}{5n^{3/2}} \]

\[ = \lim_{n \to \infty} \frac{2}{5} \sqrt{n} \]

\[ = \frac{2}{5} \cdot \infty \]

\[ = \infty \]

Chebyshev's condition does not hold.

\[ \Rightarrow \text{Cannot draw any conclusion whether WLLN holds or not.} \]

Now, \[ \lim_{n \to \infty} \frac{\psi}{\sigma^3} = \lim_{n \to \infty} \left( \frac{\sum_{k=1}^{n} k^{5/2}}{n^{13/2}} \right) \]

\[ = \lim_{n \to \infty} \left( \frac{\int_0^n x^{3/2} \, dx}{n^{13/2}} \right) \]

\[ = \lim_{n \to \infty} \left( \frac{n^{7/6}}{n^{13/2}} \right) \]

\[ = \lim_{n \to \infty} \frac{1}{n^{9/4}} \]

\[ = 0 \]
By Liapounov's CLT,
\[ S_n - \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \frac{X_k}{\sqrt{\sum_{k=1}^{n} \text{Var}(X_k)}} \xrightarrow{D} N(0,1) \]
\[ \Rightarrow S_n \xrightarrow{D} N(0, \frac{2}{5} n^{-3/2}) \]

Ex. 2: \( X, Y \) is a sequence of independent R.V.'s.
Determine if they obey WLLN and/or CLT.
\[ P[X_k = 2^k] = \frac{1}{2} = P[X_k = -2^k] \]
\[ P[X_k = 2^{-k}] = \frac{1}{2} = P[X_k = -2^{-k}] \]

Theorem: Let \( X_n \xrightarrow{P} X, Y_n \xrightarrow{P} Y \). Then
\[ aX_n \xrightarrow{P} aX, X_n + Y_n \xrightarrow{P} X + Y, X_n Y_n \xrightarrow{P} X Y, X_n \xrightarrow{P} X \text{ if } P[Y_n = 0] = 0 \Rightarrow P[Y = 0] \]
\[ f(X_n) \xrightarrow{P} f(X) \text{ if } f(\cdot) \text{ is continuous on } \mathbb{R}. \]

Proof: For \( \epsilon > 0 \),
\[ P[|X_n - x| < \epsilon] \xrightarrow{P} 1 \text{ and } P[|Y_n - y| < \epsilon] \xrightarrow{P} 1 \text{ as } n \to \infty \]
\[ P[ax_n - ax, aX_n \xrightarrow{P} aX. \]

Note that,
\[ P[|X_n - x| < \frac{\epsilon}{2}] \cap P[|Y_n - y| < \frac{\epsilon}{2}] \subseteq P[|X_n + Y_n - x - y| < \frac{\epsilon}{2}] \]
\[ \Rightarrow P[|X_n - x| > \frac{\epsilon}{2}] \cup P[|Y_n - y| > \frac{\epsilon}{2}] \supseteq P[|X_n + Y_n - x - y| > \epsilon] \]
\[ \Rightarrow P[|X_n + Y_n - x - y| > \epsilon] \leq P[|X_n - x| > \frac{\epsilon}{2}] P[|Y_n - y| > \frac{\epsilon}{2}] \xrightarrow{P} 0 \text{ as } n \to \infty \]

Hence, \( X_n + Y_n \xrightarrow{P} X + Y \).