மனோன்மணியம் சுந்தரனார் பல்கலைக்கழகம்

திருநெல்வேலி.

MANONMANIAM SUNDARANAR UNIVERSITY

TIRUNELVELI - 627 012

தொலைநிலைத் தொடர்கல்வி இயக்ககம்

DIRECTORATE OF DISTANCE AND CONTINUING EDUCATION

M.Sc. MATHEMATICS

PROBABILITY THEORY

Sub. Code: SMAM32 **(For Private Circulation only)**

PROBABILITY THEORY - SYLLABUS

Unit I. Random Events and Random Variables: Random events – Probability axioms – Combinatorial formulae – conditional probability – Bayes Theorem – Independent events – Random Variables – Distribution Function – Joint Distribution – Marginal Distribution – Conditional Distribution – Independent random variables – Functions of random variables.

Chapter 1: Sections 1.2 to 1.7 Chapter 2 : Sections 2.1 to 2.9

UNIT II. Parameters of the Distribution: Expectation Moments –The Chebyshev Inequality – Absolute moments – Order parameters –Moments of random vectors – Regression of the first and second types.

Chapter 3 : Sections 3.1 to 3.8.

UNIT III. Characteristic functions: Properties of characteristic functions – Characteristic functions and moments – semi-invariants – characteristic function of the sum of the independent random variables – Determination of distribution function by the Characteristic function – Characteristic function of multidimensional random vectors – Probability generating functions.

Chapter 4 : Sections 4.1 to 4.7.

UNIT IV. Some Probability distributions: One point , two point , Binomial – Polya – Hypergeometric – Poisson (discrete) distributions – Uniform – Normal - Gamma – Beta – Cauchy and Laplace (continuous) distributions.

Chapter 5: Section 5.1 to 5.10 (Except 5.3)

UNIT V. Limit Theorems: Stochastic convergence – Bernaulli law of large numbers – Convergence of sequence of distribution functions – de Moivre-Laplace Theorem– Lindberg Theorem – Lapunov Theroem - Poisson, Chebyshev, Khintchine Weak law of large numbers

Chapter 6: Sections 6.1 to 6.4, 6.7, 6.8 and 6.11.

Text Book: M. Fisz, Probability Theory and Mathematical Statistics, John Wileyand Sons, New York, 1963.

Contents

UNIT – I

RANDOM EVENTS AND RANDOM VARIABLES

1. Random Events

1.1.1. Preliminary

Probability theory is a part of mathematics which is useful in discovering and investigating the regular features of random events. The following examples show what is ordinarily understood by the term random event.

Example 1

Let us toss a symmetry coin the result may be either a head or a tail. We cannot predict the result. It depends various causes the initial velocity of the coin, the initial angle of through and the smoothness of the table on which the coin falls, but we cannot control all these parameters.

The result of a coin tossing head or tail is a random event.

If we perform a long series of tossing, the no.of times heads occur is approximately equal to the no.of times tails appear.

Let n denote the no.of all our tosses and m denote the no.of times heads appears.

Frequency of appearance of heads $= \frac{m}{n}$

Frequence of appearance of tails $= \frac{n-m}{n}$

Suppose we tossed a coin 4040 times and obtained heads 2048 times.

The ratio of heads $= \frac{2048}{4040} = 0.50693$

Suppose we tossed 24000 times and obtained heads 12012.

Ratio of heads $= \frac{12012}{24000} = 0.5005$

Clearly frequency oscillate about the number 0.5.

Example 2.

Let us consider the number of births of boys and girls in Poland in the year 1927 to 1932

Year in Birth	No.of Birth		Total no.of Birth	Frequency of Birth	
	Boys m	Girls f	$m+f$	Boys P_1	Girls P_2
1927	496,544	462,189	958,733	0.518	0.482
1928	513,654	477,339	990,993	0.518	0.482
1929	514,765	479,336	994,101	0.518	0.482
1930	528,072	494,739	1,022,811	0.516	0.484
1931	496,986	467,587	964,573	0.516	0.484
1932	482,431	452,232	934,663	0.516	0.484
Total	3,032,452	2,833,422	5,865,874	0.517	0.483

In this table m and f denote respectively the no.of birth of boys and girls in particular years.

$$P_1 = \frac{m}{m+f}$$

$$P_2 = \frac{f}{m+f}$$

The values of P_1 oscillate about the no. 0.517 and the values of P_2 oscillate about the no. 0.483.

Example 3.

We throw a dice. As a result of a throw one of the faces $1,2 \ldots \ldots ,6$ appears.

The appearance of any particular face is a random event.

Clearly, the frequency of this event will oscillate about the number $\frac{1}{6}$.

1.1.2. Random events and operations performed on them

We now construct the mathematical definition of a random event.

Example 1.

Suppose that when throwing a die we observe the frequency of the event, an even face.

Let elementary event $E = \{e_i / i \ = 1\ to\ 6\}$

The random event an even face will appear in e_2, e_4, e_6

$\therefore A = \{e_2, e_4, e_6\}$

Find the random event which not contain face $= 1(i.e., e_2)$

The random event $A = \{e_2, e_3, e_4, e_5, e_6\}$ which contains 5 elements.

Definition.

Every element of the Borel field $\mathbb{Z}$ of subsets of the set E of elementary events is called a ***Random event***.

Definition.

The event containing all the elements of the set E of the elementary events is called the ***sure event***.

Definition.

The event which contains no elements of the set E of elementary events is called the ***Impossible event***.

The impossible event is denoted by (0).

Definition.

We say that ***Event A is contained in event B*** if every elementary event belonging to A belongs to B.

We write $A \subset B$.

In figure 1.2.1., where square E represents the set of elementary events and circles A & B denote subsets of E. Clearly $A \subset B$.

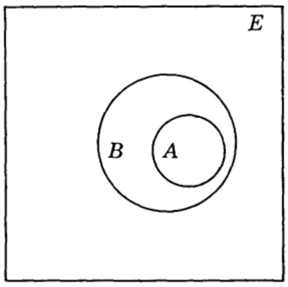


Figure 1.2.1

Definition.

Two events A & B are ***equal*** if A is contained in B & B is contained in A.

We write $A = B$.

Properties of Borel field (Z)

Property 1. The set Z of random events contains as an element the whole set E.

Property 2. The set Z of random events contains as an element the empty set (0)

i.e., Z contains sure event and impossible event.

Definition.

Two events A and B are ***exclusive*** if they do not have any common element of the set E.

Definition.

Let $A_1, A_2, A_3, \dots \dots$ be a finite or denumerable sequence of random event. The event A which contains elementary events, which belongs to atleast one of the events $A_1, A_2, \dots \dots A$ is called the ***alternative (or sum or union)*** of the events $A_1, A_2, A_3, \dots \dots$

We write,

$$A = A_1 \cup A_2 \cup A_3 \cup \ldots \ldots ..$$

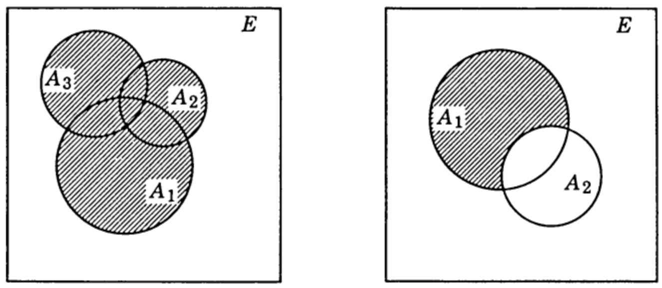
(or)

$$A = A_1 + A_2 + \cdots ....$$

(or)

$$A = \sum_{i\in} A_i.$$

Example 2.



On this figure, square E represents the set of elementary events and circles $A_1, A_2,$ ...denote three events; he shaded area represented alternative $A_1 + A_2 + A_3$.

Property 1.2.3

If a finite or denumerable no. of events A_1, A_2 and so on belonging to Z, then their alternative also belongs to Z.

Definition.

The random event A containing those and only those elementary events which belonging to A_1 but do not belong to A_2 is called the ***difference*** of the events A_1 and A_2.

We write $A = A_1 - A_2$

Property 1.2.4.

If events A_1 and A_2 belong to Z, then their difference also belongs to Z.

Example 3.

Suppose we consider the no.of children in a group of families. Let $E = \{e_0, e_1, \dots \dots, e_n\}$. Consider the event A that a family chosen at random has only one child and the event B that the family has atleast one child.

i.e., $A = \{family\ with\ one\ child\}$

$B = \{family\ with\ atleast\ one\ child\}$

Clearly, $A \subseteq B$ and

Alternative $A + B = B$

$\therefore$ E has $n + 1$ elements

Then $A = \{e_1\}$ & $B = \{e_1, e_2, \dots \dots, e_n\}$

$A - B = 0$

$B - A = \{e_2, e_3, \dots \dots, e_n\}$

i.e., $B - A$ is the event that the family has more than one child.

Definition.

The event A contains elements which belong to all the events $A_1, A_2, \dots ..$ is called the ***product (on intersection)*** of these events. We write

$A = A_1 \cap A_2 \cap \dots \dots .(or)\ A = A_1 A_1 \dots \dots ..(or) A = \prod_i A_i$.

Property 1.2.5.

If a finite (or) denumerable no. of events $A_1, A_2, \dots \dots$ belong to Z, then their product also belongs to Z.

Example 4.

Consider the random event A that a form chosen at random has atleast one horse & one plow, with the additional condition that the maximum no. of plows as were as the maximum no. of horses are two. Consider also the event B that on the form these is exactly one horse and at most one plow. Find the product of events A and B.

Solution.

$E = \{e_{00}, e_{01}, e_{02}, e_{10}, e_{11}, e_{12}, e_{20}, e_{21}, e_{22}\}$ where the 1st index denoting the no. of horses and the 2nd the no. of plows.

$A = \{e_{11}, e_{12}, e_{21}, e_{22}\}$

$B = \{e_{11}, e_{10}\}$

The product $A \cap B = \{e_{11}\}$

The event $A \cap B$ occurs iff on the chosen form there is exactly one horse and exactly one plow.

Definition

The difference of events $E = A$ is called the ***complement*** of the event A and is denoted by $\bar{A}$.

Example 5.

Suppose we have a no. of electric light bulbs. We fix a certain value to such that if the burns out in a time shorter than to, we consider it to be defective. Find the random event that we select a good bulbs.

Solution.

Consider E as all the electric light bulbs. Consider the random events A = defective bulb

i.e., $AS = \{bulb\ burns\ out\ in\ a\ time\ shorter\ than\ to\}$

$$\bar{A} = E - A$$

={bulbs that glows for a time no shorter than to}

i.e. $\bar{A}$ is a not defective

i.e., $\bar{A}$ is the random event to select good bulb.

Definition.

A set Z of subsets of the set E of elementary events with properties 1 to 5 is called a ***Borel field of events*** and its elements are called ***random events***.

i.e. (i) $E, (0)$ belongs to Z

(ii) $A_1 + A_2 + \cdots \ldots \in Z$

(iii) $A_1 \cap A_2 \cap \ldots \ldots \cap \in Z$

(iv) $A_1 - A_2 - \cdots \ldots \in Z$

Definition.

The sequence $\{A_n\}$ $(n = 1,2, \ldots \ldots)$ of events is called ***non-increasing*** if for every n we have $A_n \supset A_{n+1}$.

Definition.

The product of a non-increasing sequence of events $\{A_n\}$ is called the ***limit*** of this sequence.

Write $A = \prod_{n\geq 1} A_n = \lim_{n\to\infty} A_n$.

Definition.

The sequence $\{A_n\}(n = 1,2, \ldots \ldots)$ of events is called ***non-decreasing*** if for every n we have

$A_{n+1} \supset A_n$

The sum of a non-decreasing sequence $\{A_n\}$ is called the ***limit*** of this sequence.

We write,

$$A = \sum_{n\geq 1} A_n = \lim_{n\to\infty} A_n$$

1.1.3. The system of axioms of the theory of probability

Axiom I: To every random event A there corresponds a certain number $P(A)$ called the probability of A, which satisfies the inequality $0 \leq P(A) \leq 1$.

Example 1.

Suppose there are only black balls in an urn. Let the random experiment consist in drawing a ball from urn. Drawing the black ball out of the urn is a sure event.

Axiom II: The probability of sure event equals one.

i.e., $P(E) = 1$.

Example 2.

Find the frequency of face 6 and face 2 in a dice.

Solution.

Let A = getting face

B = getting face 2

$\therefore P(A) = \frac{1}{6}$ and $P(B) = \frac{1}{6}$

Probability of getting either face 6 or face 2 $= P(A) + P(B)$

$$= \frac{1}{6} + \frac{1}{6}$$

$$= \frac{2}{6}$$

$$= \frac{1}{3}$$

$\therefore$ The frequence of getting either face 6 or face 2 $= \frac{1}{3}$

Example 3.

If a card is selected from a deck of 52 cards many times over. Find the frequency of

(i) Appearance of ace

(ii) Appearance of spade

(iii) Appearance of ace or spade

Solution.

Let A be the event of getting ace

Let B be the event of getting spade

(i) Let $P(A) = probability\ of\ getting\ ace$

$$= \frac{4}{52}$$

(ii) Let $P(B) = probabilit\ of\ getting\ spade$

$$= \frac{13}{52}$$

(iii) Let $P(A \cup B) = probability\ of\ getting\ ace\ or\ spade$

$$\therefore P(A \cap B) = \frac{1}{52}$$

$$\therefore P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$= \frac{4}{52} + \frac{13}{52} - \frac{1}{52}$$

$$= \frac{4+13-1}{52}$$

$$= \frac{16}{52} = \frac{8}{26} = \frac{4}{13}.$$

Axiom III: The probability of the alternative of a finite or denumerable no. of pairwise exclusive events equals the sum of the probabilities of these events.

Thus, if We have a finite or countable sequence of pairwise exclusive events $\{A_k\}, k = 1,2,3 \dots \dots$, then axiom 3 the following formula holds:

$P(\sum_k A_k) = \sum_k P(A_k)$.

In particular, if the random event contains a finite or countable number of elementary events e_k and $e_k \in Z(k = 1,2, \dots)$

$P(e_1, e_2, \dots \dots) = P(e_1) + P(e_2) + \cdots \dots . \forall e_K \in Z, K = 1,2,3, \dots \dots$

Note.

1. Axiom 3 is called the countable (or complete) additivity of probability.
2. P is said to be probability if it satisfying axiom 1, axiom 2, axiom 3.
3. P(A) satisfying Axiom 1,2,3 is normal, non-negative and countably additive measure on the Boret field Z of subsets of E

Theorem 1.1.

Let A and B be two arbitrary random events, exclusive or not. Then $P(A \cup B) = P(A) + P(B) = P(A \cap B)$

Proof.

The set $A \cup B$ and B can be written $A \cup B = A \cup (B - AB) \rightarrow (1)$ and $B = AB \cup (B - AB) \quad \rightarrow (2)$

Here, in (1) & (2)

A and $(B - AB), B$ and $(B - AB)$ are exclusive.

∴ By Axiom (3)

(1) $\Rightarrow P(A \cup B) = P(A) + P(B - AB) \rightarrow (3)$

$P(B) = P(AB) + P(B - AB)$

$P(B - AB) = P(B) - P(AB) \rightarrow (4)$

Sub (4) in (3),

$P(A \cup B) = P(A) + P(B) - P(AB)$

Remark.

Let $A_1, A_2, \ldots\ldots, A_n$ where $n \geq 3$, be arbitrary random events

$$P(\sum_{k=1}^{n} A_K) = \sum_{k=1}^{n} P(A_K) - \sum_{\substack{k_1,k_2=1 \\ k_1<k_2}}^{n} P\left(A_{k_1} \cap A_{k_2}\right) + \sum_{\substack{k_1,k_2,k_3=1 \\ k_1<k_2<k_3}}^{n} P\left(A_{k_1} \cap A_{k_2} \cap A_{k_3}\right) - \cdots \ldots\ldots + (-1)^{n+1} P\left(A_{k_1} A_{k_2} \ldots\ldots\ldots A_n\right)$$

Theorem 1.2.

If the events $A_1, A_2, \ldots\ldots$ exhaust the set of elementary events $E, P\left(\sum_{k=1} A_K\right) = 1$

Example 4.

Let the set of all non-negative integers form the set of elementary events. Let (e_n) be the event of obtaining the number n where $n = 0,1,2, \ldots\ldots$ suppose that $(e_n) = \frac{C}{n!}$, where C is a constant. Prove that $C = e^{-1}$.

Solution.

Let $E = Z^+ \cup \{0\}$

$Z = \{e_0, e_1, e_2 \ldots\ldots\}$ where (e_i) be the event of obtaining the number i where $i = 0,1,2,3, \ldots\ldots$

Given $P(e_n) = \frac{C}{n!}$

$$P(\sum_{n=0}^{\infty} e_n) = \sum_{n=0}^{\infty} P(e_n)$$

$$= \sum_{n=0}^{\infty} \frac{C}{n!}$$

$$= C \sum_{n=0}^{\infty} \frac{1}{n!}$$

We know that,

$$P(\sum_{n=0}^{\infty} e_n) = 1 \ \& \ \sum_{n=0}^{\infty} \frac{1}{n!} = e$$

$$P(\sum_{n=0}^{\infty} e_n) = C \sum_{n=0}^{\infty} \frac{1}{n!}$$

$\Rightarrow 1 = ce$

$c = \frac{1}{e}$

$\therefore C = e^{-1}.$

Theorem 1.3.

The probability of the impossible event is zero.

Proof.

For every random event A, we have

$$A \cup E = E$$

If A is the impossible event $A = (0)$

Then, A and E are exclusive

From axiom (3)

$$P(A \cup E) = P(A) + P(E)$$

$$P(E) = P(A) + P(E)$$

$$P(A) = P(E) - P(E)$$

$$P(A) = 0$$

$$\therefore P\big((0)\big) = 0$$

Theorem 1.4.

Let $\{A_n\}; n = 1,2,3, \dots \dots$ be a non-increasing sequence of events and let A be their product. Then $P(A) = \lim_{n\to\infty} P(A_n)$.

Proof.

If the sequence $\{A_n\}$ is non- increasing,then for every n we have

$$A_n = \sum_{k=n}^{\infty} A_K \overline{A_{K+1}} + A$$

sWe know that,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$\mathrm{P}(A_n) = P(\sum_{k=n}^{\infty} A_K \overline{A_{K+1}}) + P(A) - P\big((\sum_{k=n}^{\infty} A_K \overline{A_{K+1}})A\big)$$

$$\mathrm{P}(A_n) = P(\sum_{k=n}^{\infty} A_K \overline{A_{K+1}}) + P(A) - P(\sum_{k=n}^{\infty} A_K \overline{A_{K+1}}\, A) \;\rightarrow (1)$$

By axiom (III),

$P(\sum_{k=n}^{\infty} AA_K \overline{A_{K+1}}) = 0$ ($\because \forall k$ the event $AA_K\overline{A_{K+1}}$ is the impossible event)

$(1) \Rightarrow P(A_n) = P(\sum_{k=n}^{\infty} A_K \overline{A_{K+1}}) + P(A)$

$\mathrm{P}(A_n) = \sum_{k=n}^{\infty} P(A_K \overline{A_{K+1}}) + P(A)$

Taking limit $n \rightarrow \infty$ on both sides.

$$\lim_{n\to\infty} P(A_n) = \lim_{n\to\infty} (\sum_{k=n}^{\infty} P(A_K \overline{A_{K+1}}) + P(A)\,)$$

The series $\sum_{k=1}^{\infty} P(A_K \overline{A_{K+1}})$ is convergent being a sum of non-negative terms whose partial sums are bounded by one.

$$\therefore \lim_{n\to\infty} P(A_n) = P(A)$$

Theorem 1.5.

Let $\{A_n\}, n = 1,2,3, \dots \dots$ be a non-decreasing sequence of the events and let A be the alternative, then we have $P(A) = \lim_{n\to\infty} P(A_n)$

Proof.

Consider, the sequence of the events $\{\overline{A_n}\}$ which are the complements of the event A_n.

By our assumption thtat $\{A_n\}$ is non-decreasing

$\Rightarrow \{\overline{A_n}\}$ is non-increasing sequence

Let $\bar{A}$ be the product of events $\overline{A_n}, \forall n$

Then by the theorem 1.3.4,

$$P(\bar{A}) = \lim_{n\to\infty} P(\overline{A_n}) \quad \rightarrow (1)$$

Thus, $P(A) = 1 - P(\bar{A})$

$$= 1 - \lim_{n\to\infty} P(\overline{A_n}) \quad (From\ (1))$$

$$= 1 - \lim_{n\to\infty} P(1 - A_n)$$

$$= 1 - \lim_{n\to\infty} [P(1) - P(A_n)\,]$$

$$= 1 - 1 + \lim_{n\to\infty} P(A_n)$$

$\therefore P(A) = \lim_{n\to\infty} P(A_n)$

Theorem 1.6.

If the events A and B satisfy the condition $A \subset B$ then $P(A) \leq P(B)$.

Proof.

Given, $A \subset B$

$$B = A + (B - A)$$

Events A & $B - A$ are exclusive

$\therefore \quad P(B) = P(A + (B - A))$

$$= P(A) + P(B - A) \quad (by\ axiom\ 3)$$

Since $P(B - A) \geq 0$ we have $P(B) \geq P(A)$.

1.1.4. Application of combinatorial formulas for computing probabilities

In some problems we can compute probabilities by applying combinatorial formulas. we illustrate this by some examples.

Example 1.

Suppose we have 5 balls of different colors in an urn. Assume that the probabilities of drawing any particular ball is the same for any ball and equal p.

Here E consists of 5 elements and by hypothesis each has the same probability.

Hence by theorem 1.3.1, we have $5p = 1, or\ p = \frac{1}{5}$.

Example 2.

Suppose we have in the urn 9 slips of papers with the numbers 1 to 9 written on them, and suppose there are no two slips marked with the same number. Then E has 9 elementary events. Denote by A the event that on the slip of paper selected at random an even number will appear. What is the probabilities of this event?

Solution.

Suppose that the probability of selecting any particular slip is the same for any slip, and hence equals it. We shall obtain a slip with an even number if we draw one of the slips marked with 2, 4, 6 or 8.

According to axiom III, the required probability equals

$$P(A) = \frac{1}{9} + \frac{1}{9} + \frac{1}{9} + \frac{1}{9} = \frac{4}{9}$$

If we compute the probability of selecting a slip with an odd number, we may notice that this random event is the complement of A (we denote it by A) and, by theorem 1.3.2, we have

$$P(\bar{A}) = 1 - P(A) = \frac{5}{9}.$$

Example 3.

Let us toss a coin three times. What is the probability that heads appear twice?

Solution.

The number of all possible combinations which may occur as a result of three successive tosses equals $2^3 = 8$. We have the following possible combinations:

$$HHH, HHT, HTH, THH, HTT, THT, TTH, TTT$$

Consider each of these combinations as an elementary event and the whole collection of them as the set E. Suppose that the occurrence of each of them has the same probability. Then We have that the probability of each particular combination equals $\frac{l}{2^3}$. From the table we see that heads appear twice in three elementary events

$$E_1 = \{HHT, HTH, THH\}$$

Hence by axiom III the required probability is $P(E_1) = \frac{3}{8}$.

Example 4.

If we toss a coin n times. What is the probability that heads appear twice?

Solution.

Here we toss a coin n time. The number of all possible combinations with n tosses equals 2n. The number of combinations in which heads appear in times equals the number of combinations of in elements from n elements given by

$$\binom{n}{m} = \frac{n!}{m!\,(n-m)!}$$

If every possible result of n successive tosses of a coin is equally likely, the required probability is

$$\frac{n!}{2^n m!\,(n-m)!}$$

Example 5.

Compute the probability that heads appear at least twice in three successive tosses of a coin.

Solution.

The random event under consideration will occur if in three tosses heads appear two or three times.

According to formula in the result, the probability that heads appear three times equals

$$\frac{3!}{2^3.\,3!\,.0!} = \frac{1}{8}$$

and the probability that heads appear twice equals $\frac{3}{8}$.

Hence, according to axiom III, the required probability is

$\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$

Note. In examples 1 to 4 the equiprobability of all elementary events was assumed. This assumption was obviously satisfied in our examples, but it is not always acceptable.

1.1.5. Conditional Probability

Definition.

Let the probability of the event B be positive. The ***conditional probability of the event A provided B has occurred*** equals the probability of AB divided by the probability of B. Thus

$$P(A|B) = \frac{P(AB)}{P(B)}, where\ P(B) > 0 \ldots\ldots (1)$$

$$P(B|A) = \frac{P(AB)}{P(A)},\ where\ P(A) > 0 \ldots\ldots (2)$$

From (1) and (2), we obtain,

$$P(AB) = P(B)P(A|B) = P(A)P(B|A) \ldots\ldots (3)$$

This formula is to be read: *The probability of the product AB of two events equals the product of the probability of B times the conditional probability of A provided B as occur or the product of the probability of A times the conditional probability of B provided A as occurred.*

Let A_1, A_2, A_3 denoted three events from the same field Z.

The probability of A_3 provided the product A_1A_2 has occurred equals

$$P(A_3|A_1A_2) = \frac{P(A_3A_1A_2)}{P(A_1A_2)} \text{ where } P(A_1A_2) > 0 \ldots\ldots (4)$$

From (1) and (4) we obtain the for the probability of the product of three events the relations

$$P(A_1A_2A_3) = P(A_1A_2)P(A_3|A_1A_2)$$

$$= P(A_1)P(A_2|A_1)P(A_3|A_1A_2)$$

This formula is to be read: *The probability of the product of three events equals the probability of the first event times the conditional proba-bility of the second event provided the first event* *has occurred times the probability of the third event provided the product of the first two events has occurred.*

Let $A_1, A_2, \dots \dots A_n$ be random events. Consider the conditional probabilities $P(A_{k_1} A_{k_2} \dots A_{k_r} | A\_k_{r+1} \dots A_{k_n})$ of the product of some subgroup consisting of r events $(1 \leq r \leq n-1)$ provided the product of the remaining $n - r$ events has occurred. Then we obtain

$$P(A_1 A_2 \dots \dots A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 A_2) \dots \dots P(A_n|A_1 A_2 \dots \dots A_{n-1}).$$

Remark. We shall show that the conditional probability satisfies axioms I to III.

We know that $P(AB) \leq P(B)$

Event Bmay occur either when event A occurs or when event A does not occur. hence

$$B = AB \cup \bar{A}B,$$

where $\bar{A}$ is complement of A. Thus

$$AB \subset AB \cup \bar{A}B \; P(AB) = P(B)P(A|B)$$

Since $P(AB) \geq 0, P(B) > 0$ we obtain

$$0 \leq P(A|B) \leq 1$$

Which is the property expressed by axiom I.

Now, let $A|B$ be the sure event in Z'. That is let $AB = B$. Then

$$P(AB) = P(B)$$

And hence

$$P(A|B) = \frac{P(AB)}{P(B)} = 1$$

This is the property expressed by axiom II.

Consider the alternative $\sum_i (A_i \,|B)$ of pairwise exclusive events. We can write

$$\sum_i (A_i \,|B) = (\sum_i A_i)|B ,$$

And hence

$$P[\sum_i (A_i\,|B)] = P[(\sum_i A_i)|B\,]$$

According to (1) and the axiom III we have

$$P[\sum_i (A_i\,|B)] = \frac{P[(\sum_i A_i)B]}{P(B)}$$

$$= \frac{P(\sum_i A_i\,B)}{P(B)}$$

$$= \sum_i \frac{P(A_i B)}{P(B)}$$

$$= \sum_i P(A_i|B)$$

$\therefore P[\sum_i (A_i|B)] = \sum_i P(A_i\,|B)$ s

This formula expresses the countable additivity of the conditional probability.

Since the axioms are satisfied for the conditional probabilities, the theorems derived from these axioms hold for the conditional probabilities.

1.1.6. Bayes Theorem

Let us consider the following examples.

Example 1.

We have 2 urns. There are 3 white and 2 black balls in the 1^{st} urn and 1 white & 4 black balls in the second urn. From an urn chosen at random we select one ball at random. What is the probability of obtaining a white ball if the probability of selecting each of the urns equals 0.5?

Solution.

Let A_1 & A_2 be the events of selecting the 1^{st} or 2^{nd} urn respectively

Given the probability of selecting each of urns equals 0.5

i.e., $P(A_1) = 0.5$ & $P(A_2) = 0.5$

let B be the event of selecting the white ball.

The probability of selecting white ball from A_1 is $\frac{3}{5} = 0.6$

i.e., $P(B|A_1) = 0.6$

The probability of selecting white ball from A_2 $is \frac{1}{5} = 0.2$

i.e., $P(B|\backslash A_2) = 0.2$

Since B is the event of selecting white balls.

$$B = A_1B + A_2B$$

Since events A_1B and A_2B are exclusive.

$$P(B) = P(A_1B) + P(A_2B)$$

We know that, $P(AB) = P(A)P(B|A)$

$$\therefore P(B) = P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$$

$$= (0.5)(0.6) + (0.5)(0.2)$$

$$= 0.30 + 0.10$$

$$= 0.40$$

$\therefore$ The probability of obtaining white balls = 0.4.

Theorem 1.7.(Theorem of absolute probability)

If the random events $A_1, A_2, \ldots$ re pairwise exclusive and exhaust the set E of elementary events and if $P(A_i) > 0$ for $i = 1,2,3, \ldots$.Then for any random event B we have

$$P(B) = P(A_1)P(B\backslash A_1) + P(A_2)P(B\backslash A_2) + \cdots \text{....}$$

Proof.

Let B be any random event.

Since $A_1, A_2, \ldots$ are pairwise exclusive.

$\therefore$ B may together with one and only one of events A_i. Then we have

$$B = A_1B + A_2B + \cdots$$

$$P(B) = P(A_1B) + P(A_2B) + \cdots$$

We know that $P(AB) = P(A)P(B|A)$

∴ We have,

$$P(B) = P(A_1)P(B\backslash A_1) + P(A_2)P(B\backslash A_2) + \cdots ....$$

Hence the proof.

Theorem 1.8. (Bayes theorem)

If the random events $A_1, A_2, \dots \dots$ are pairwise exclusive and exhaust the set E of elementary events and if $P(A_i) > 0$ for $i = 1,2,3, \dots$ we have

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+\cdots}$$

Proof.

We know that $P(A|B) = \frac{P(AB)}{P(B)}$

Substitute A by A_i,

$$P(A_i\backslash B) = \frac{P(A_iB)}{P(B)} = \frac{P(A_i)P(B\backslash A_i)}{P(B)}$$

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+\cdots ...}$$

Example 2.

Guns I and 2 are shooting at the same target. It has been found that gun 1 shoots on the average nine shots during the same time gun 2 shoots ten shots. The precision of these two guns is not the same; on the average, out of ten shots from gun 1 eight hit the target, and from gun 2, only seven. During the shooting the target has been hit by a bullet, but it is not known which gun shot this bullet. What is the probability that the target was hit by gun 2?

Solution.

Denote by Al and A2 the events that a bullet is shot by gun 1 and gun 2, respectively. Taking into consideration the ratio of the average number of shots made by gun 1 to the average number of shots made by gun 2, we can put $P(A_1) = 0.9P(A_2)$.

Denote by B the event that the target is hit by the bullet. According to the data about the precision of the guns we have

$$P(B|A_1) = 0.8 \; and \; P(B|A_2) = 0.7$$

According to Bayes formula

$$P(A_2|B) = \frac{P(A_2)P(B|A_2)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}$$

$$= \frac{0.7\, P(A_2)}{0.9\, P(A_2)\, 0.8 + 0.7\, P(A_2)} = 0.493$$

Exercise.

1. A deck of cards contains 52 cards. Player G has been dealt 13 of them. Compute the probability that player G has

(a) Exactly 3 aces
(b) At least 3 aces
(c) Any 3 face cards of the same face value
(d) Any 3 cards of the same face value from the 5 highest denominations
(e) Any 3 cards of the same face value from the eight lowest denominations
(f) Any 3 cards of the same value,
(g) Three successive spades
(h) At least three successive of any suit
(i) Three successive cards of any suit
(j) At least three successive cards of any suit

1.1.7. Independent Events

In general, the conditional probability $P(A|B)$ differs from $P(A)$.

Suppose $P(A|B) = P(A)$

We know that $P(AB) = P(B)P(A|B)$

$\Rightarrow P(AB) = P(B)P(A)$

Suppose $P(B|A) = P(B)$

$\therefore P(AB) = P(A)P(B|A)$

$\Rightarrow P(AB) = P(A)P(B)$, Where $P(A) > 0$ and $P(B) > 0$.

Definition.

Two events A and B are called ***independent*** if $P(AB) = P(A)P(B)$, where $P(A) > 0$ & $P(B) > 0$.

i.e., If the probability of the product AB is equal to the product of the probabilities A and B.

Definition.

Events $A_1, A_2, \dots \dots, A_n$ are ***independent*** if for all integers indices $k_1, k_2, \dots, k_s$ satisfying the conditions

$$1 \leq K_1 < K_2 < \cdots \dots < K_s \leq n$$

We have

$$P\left(A_{k_1}, A_{k_2} \dots \dots A_{k_s}\right) = P\left(A_{k_1}\right)P\left(A_{k_2}\right) \dots \dots . P\left(A_{k_s}\right)$$

i.e., If the probability of the product of every combination $A_{k_1}, A_{k_2}, \dots \dots . A_{k_s}$ of events equals the product of the probabilities of these events.

Example 1.

There are 4 slips of paper of identical size in an urn. Each slip is marked with one of the numbers 110, 101, 011, 000 and there are no two slips marked with the same number. Consider event A_1 that on the slip selected the number – 1 appears in the first place, event A_2 that one appears in the second place and A_3 that one appears in the third place. Verify A_1, A_2, A_3 are independent.

Solution.

Let $A_1 = \{110,101\}$

$P(A_1) = \frac{2}{4} = \frac{1}{2}$

Let $A_2 = \{110,011\}$

$P(A_2) = \frac{1}{2}$

Let $A_3 = \{101,011\}$

$P(A_3) = \frac{1}{2}$

$A_1A_2A_3 = \{\emptyset\}$

$P(A_1A_2A_3) = 0$

$P(A_1)P(A_2)P(A_3) = \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} = \frac{1}{8} \neq 0$

$\therefore P(A_1A_2A_3) \neq P(A_1)P(A_2)P(A_3)$

$\therefore A_1A_2A_3$ are not independent

$A_1A_2 = \{110\}$

$P(A_1A_2) = \frac{1}{4}$

$A_2A_3 = \{011\}$

$P(A_2A_3) = \frac{1}{4}$

$A_1A_3 = \{101\}$

$P(A_1A_3) = \frac{1}{4}$

$P(A_1) = \frac{1}{2}; P(A_2) = \frac{1}{2}; P(A_3) = \frac{1}{2}$

$P(A_1)P(A_2) = \frac{1}{2}.\frac{1}{2} = \frac{1}{4} = P(A_1A_2)$

$P(A_2)P(A_3) = \frac{1}{4} = P(A_2A_3)$

$P(A_1)P(A_3) = \frac{1}{4} = P(A_1A_3)$

$\therefore A_1, A_2, A_3$ are independent.

Definition.

Events $A_1, A_2, \ldots$ are ***independent*** if for every $n = 2,3 \ldots \ldots$ events $A_1, A_2, \ldots \ldots, A_n$ are independent.

Example 2.

Pairwise independent does not imply mutual independence. Suppose we twice spin a fair spinner with the numbers 1,2,3 & 4. Let A_1 be the event that sum of the numbers spun is 5. Let A_2 be the event that the first number spun is a one. Let A_3 be the event the second number spun is a four. Prove that A_1, A_2, A_3 are pairwise independent but not that mutually independent.

Solution.

$P(A_1) = \frac{1}{4}\left[\frac{4}{16} = \frac{1}{4}\right]$

$P(A_2) = \frac{1}{4}$

$P(A_3) = \frac{1}{4}$

$P(A_1A_2) = \frac{1}{16} = \frac{1}{4}.\frac{1}{4} = P(A_1)P(A_2)$

$P(A_1A_3) = \frac{1}{16} = \frac{1}{4}.\frac{1}{4} = P(A_1)P(A_3)$

$P(A_2A_3) = \frac{1}{16} = P(A_2)P(A_3)$

$\therefore A_1, A_2, A_3$ are pairwise independent.

$A_1A_2A_3 = \{(1,4)\}$

$P(A_1A_2A_3) = \frac{1}{16}$

$P(A_1)P(A_2)P(A_3) = \frac{1}{4} \times \frac{1}{4} \times \frac{1}{4}$

$= \frac{1}{64}$

$\therefore P(A_1A_2A_3) \neq P(A_1)P(A_2)P(A_3)$

$\therefore A_1, A_2, A_3$ are not mutually independent.

Exercise.

1. Prove that if the events A and B are independent, the same is true for the events $\bar{A}$ and $\bar{B}$.

1.2. Random variables

1.2.1. The concept of random variable

We can assign a number to every elementary event from a set E of elementary events. In the coin-tossing example we assigned the number 1 to the appearance of heads and the number 0 to the appearance of tails. Then the probability of obtaining the number 1 as a result of an experiment will be the same as the probability of obtaining a head, and the probability of obtaining the number 0 will be the same as the probability of obtaining a tail.

Definition.

Let $X(e)$ be a single-valued seat function defined on the set E of elementary events. The set A of all elementary events to which the function $X(e)$ assigns values in a given set S of real number is called the ***inverse image of the set*** S.

Clearly the inverse image of the set $\mathbb{R}$ of all real numbers is the whole set E.

Definition.

A single-valued real function $X(e)$ defined on the set E of elementary events is called a ***random variable*** if the inverse image of every interval I in the real axis of the form $(-\infty, x)$ is a random event.

Note.

1. We can write X instead of $X(e)$
2. Random variables are usually denoted by capital letters X, Y and so on and their values by corresponding small letters $x, y, \dots \dots$

Definition.

The function $P^{(x)}(S)$ giving the probability that a random variable X takes on a value belonging to S, where S is an arbitrary borel set on the real axis, is called the ***probability function*** of X

We write,

$$P^{(x)}(S) = P^{(x)}(X \in S)$$

Remark.

1. The probability $P^{(x)}(I)$ that the random variable $X(e)$ takes on the values in the interval I equals to the probability$P(A)$ of the inverse image A of I
2. If a random event A is the inverse image of a point x, the probability that the random variable X takes on the value x equals the probability of the event A
   i.e., $P^{(x)}(X = x) = P(A)$
3. Since any interval I of the form $[a, b)$ where $a < b$, is the difference of the intervals $(-\infty, b) - (-\infty, a)$

1.2.2. The distribution function

It is convenient to characterize the probability distribution by means of the distribution function which is now defined.

Example 1.

Consider tossings of a die. To every elementary event, that is, to every result of a throw, we can assign one of the numbers 1,2,3,……,6, the number of dots which appear on the resultant face. Find

i) $P(X < 1)$

ii) $P(X < x)$ if $1 < x \leq 2$

iii) $P(X < x)$ if $2 < x \leq 3$

iv) $P(X < x)\ \ if\ 5 < x \leq 6$

v) $P(X < x)\ \ if\ x > 6$

Solution.

Let X be the random variable it takes six values. $x_i = i(i = 1,2,3,4,5,6)$

Clearly,

$$P(X = x_i) = \frac{1}{6}, for\ i = 1\ to\ 6$$

(i) $P(X < 1) = 0$

(ii) If $1 < x \leq 2$

$$P(X < x) = P(X < 2)$$
$$= P(X = 0) + P(X = 1)$$
$$= 0 + \frac{1}{6} = \frac{1}{6}$$

(iii) If $2 < x \leq 3$

$$P(X < 3) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= 0 + \frac{1}{6} + \frac{1}{6}$$
$$= \frac{2}{6} = \frac{1}{3}$$

(iv) If $5 < x \leq 6$

$$P(X < x) = P(X < 6)$$
$$= \sum_{i=1}^{5} P(X = i)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{5}{6}$$

(v) If $x > 6$

$$P(X < x) = P(X \leq 6)$$
$$= \sum_{i=1}^{6} P(X = i)$$
$$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{6}{6} = 1$$

Remark.

We obtain the step function for the above example x increases the value of $P(X < x)$ is increasing by a constant number $P(X = x_i)$.

Definition.

The function $F(x)$ is defined as $F(x) = P(X < x)$ is called the ***distribution function*** of the random variable X.

Theorem 1.9.

The single-valued function $F(x)$ is a distribution function iff it is non-decreasing, continuous at least from the left and satisfies the condition $F(-\infty) = 0, F(\infty) = 1$

Proof.

Suppose the single-valued function $F(x)$ is a distribution.

Clearly $F(-\infty) = 0, F(+\infty) = 1$

Claim: $F(x)$ is the non – decreasing function

Let x_1& x_2 where $x_1 < x_2$ be two point on the real axis.

Since $(-\infty, x_2)$ contains the interval $(-\infty, x_1)$

i.e., $(-\infty, x_1) \leq (-\infty, x_2)$

$$P(X < x_1) \leq P(X < x_2)$$

$$F(x_1) \leq F(x_2)$$

$\therefore$ $F(x)$ is a non-decreasing.

Claim: Every distribution function is continuous at least from left.

Let $x_1 < x_2 < \cdots \ldots < x$ be an arbitrary increasing sequence converge to x.

Let A_k be the event that the random variable X takes on a value from the half open interval $[x_k, x)$

If $k_1 < k_2$, from the occurrence of the event A_{k_2} follows the occurrence of the event A_{k_1}

$\therefore$ $\{A_k\}$ is the non-increasing sequence of the events.

The limit of the sequence $\{x_k\}$ is the point x, does not belong to any of the intervals.

i.e., $x \in [x_k, x)$

$\therefore$ The product $A = \prod_{k=1}^{\infty} A_K$ is the impossible

$\Rightarrow$ A is impossible event.

$\Rightarrow$ $P(A) = 0.$

By theorem 1.3.4,

$$\lim_{k\to\infty} P(A_k) = P(A)$$

i.e., $\lim_{k\to\infty} P(A_K) = P(A)$

$\therefore \lim_{k\to\infty} P(A_k) = 0$

$$\lim_{k\to\infty} P(x_k \le X < x) = 0$$

$$\lim_{k\to\infty} [F(x) - F(x_k)] = 0$$

$$\lim_{k\to\infty} F(x_k) = F(x)$$

$\therefore F(x)$ is the continuous function from the left.

Conversely, Suppose $F(x)$ is a non-decreasing and continuous from the left & $F(-\infty) = 0, F(+\infty) = 1$.

Claim: $F(x)$ is a distribution function.

Take the interval $[0,1]$ as the set of elementary events.

The field of all Borel subsets of this interval as the field of random events.

Take as a probability measure the Lebesgue measure $[i.e., P(A) = m(A)]$.

Then the probability of a Borel set from the interval $[0,1]$ is equal to its Lebesgue measure.

$i.e, A \in [0,1] \Rightarrow P(A) = m(A)$

In particular, the probability of the interval $[0, e]$ where $0 < e \le 1$ equals the length e of this interval

$\big(i.e., P([0, e]) = l([0, e] = 2)\big)$

Now we define the random variable $X(e)$ in the following way:

i.e., $X(e) = \inf_{F(y)=e} y (0 \le e \le 1)$

Thus, for a given value e, the random variable $X(e)$ equal to l.u.b of the set of all y such that $F(y) = e$.

Since $P(X|e) < Z = P\left(\inf_{F(y)=e} y < x\right)$

$$= F(x) \quad (-\infty < x < \infty)$$

$$\therefore P(X(e) < x) = F(x)$$

Thus, the distribution function of $X(e)$ is the function F.

Hence proved.

Remark.

The set of points of dis continuity is at most countable. The set of points at which the distribution function $F(x)$ has a jump not smaller than $\frac{1}{n}$ is denoted by H_n.

Then $H = H_1 + H_2 + \cdots$

For every n the set H_n is finite, hence the set H is at most countable.

1.2.3. Random Variable of The Discrete Type and The Continuous Type

Definition.

A random variable is said to be of the ***discrete type*** if it takes on, with a probability 1, values belonging is a set S which is at most countable and every value in the set S has the positive probability.

These values are called ***jump points*** and their probabilities ***jumps***.

Example 1.

A stock of fruits contains good & defective items.

Here, two elementary events

i.e., good item (or) defective item

Let the probability of drawing good item denoted by P

Suppose, $0 < P < 1$

Let the drawing of a good item is denoted by 1 and defective item the number 0.

$\therefore$ We get the random variable of discrete type which has only two values with positive probability 1 & 0 with the probabilities P and $1 - P$ respectively.

Remark.

Let a random variable X of the discrete type take on the values $x_i(i = 1,2,3 \ldots \ldots)$ with probabilities P_i.

By the definition,

$\sum_{i=1}^{n} P_i = 1$ if the no. of jump points x_i is finite.

$\sum_{i=1}^{\infty} P_i = 1$ if no. of jump points x_i is countable.

The above definition formulated as follows:

Definition.

Let $x_i(i = 1,2,3, \ldots)$ be an arbitrary jump point of a random variable X of the discrete type. The probability that the random variable X takes on the values x_i is called the ***probability function*** of the discrete-type random variable X and we write

i.e., $P(X = x_i) = P_i$,

where the numbers $P_i(i = 1,2 \ldots \ldots)$ satisfy either $\sum_{i=1}^{n} P_i = 1$ (or) $\sum_{i=1}^{\infty} P_i = 1$.

The distributive function $F(x)$ as the form

$F(x) = \sum_{x_i<x} p_i$,

where the summation is extended over all the points x_i for which $x_i < x$.

Suppose a random variable X which has no jump points. The distribution function of such a random variable is a continuous function.

Definition.

A random variable X is said to be of the ***continuous type*** if there exists a non-negative function $f(x)$ such that for every real number x the following relation holds:

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

where $F(x)$ is the distribution function of X. The function $f(x)$ is called ***probability density*** of the random variable X.

Note. "Probability density" is also called as "density function" (or) "density".

Properties of a distribution function.

1. Every density function $f(x)$ satisfies the relation $F(+\infty) = \int_{-\infty}^{\infty} f(x)dx = 1$
2. For every real $a \ and \ b$, where $a < b$, we have

$P(a \leq X \leq b) = P(X \leq b) - P(X \leq a)$

$= F(b) - F(a)$

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

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

$\therefore P(a \leq X \leq b) = \int_{a}^{b} f(x)dx$

3. Clearly for every Borel set S, we have $P(S) = \int_{S} f(x)dx$
4. If the function $f(x)$ is continuous at some point x,

$$F'(x) = f(x)$$

Thus the continuity points of the function $f(x)$ we have

$$f(x) = \lim_{\Delta x \to 0} \frac{F(x+\Delta x) - F(x)}{\Delta x}$$

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

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

$\therefore$ Every real function $f(x)$ which is non-negative integrable over the whole real axis and satisfies the condition $\int_{-\infty}^{\infty} f(x)dx = 1$ is the probability density of a random variable X of the continuous type.

Clearly, we have the function $F(x)$ defined by the formula

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

Has all the properties of a distribution function.

Example 2.

On the set of all real numbers, define the density function $f(x)$ in the following way:

$$f(x) = \begin{cases} 0 & for\ x < 0 \\ \frac{1}{2}x & for\ 0 \leq x \leq 2 \\ 0 & for\ x > 2 \end{cases}$$

Find the distribution function $F(x)$.

Solution.

We have$F(x) = \int_{-\infty}^{x} f(x)dx$.

For $x < 0$,

$$F(x) = 0$$

For $0 \leq x \leq 2$,

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

$$= \int_{-\infty}^{0} 0dx + \int_{0}^{2} \frac{1}{2}x\, dx = \frac{x^2}{4}$$

For $x > 2$,

$$F(x) = \int_{-\infty}^{0} f(x)dx + \int_{0}^{2} f(x)dx + \int_{2}^{x} f(x)dx$$

$$= 0 + \int_{0}^{2} \frac{x}{2} dx + 0$$

$$= \left[\frac{x^2}{4}\right]_0^2 = \frac{4}{4} = 1$$

$$\therefore F(x) = \begin{cases} 0 & for\ x < 0 \\ \frac{x^2}{4} & for\ 0 \leq x \leq 2 \\ 1 & for\ x > 2 \end{cases}$$

Remark.

1. If for a random variable X of a continuous type, the probability of a certain event equals 0. It does not follow that this event is impossible. It should be considered only as an event which is very unlikely is occur
2. If for a random variable X of the its type, the probability of a certain event equals 1 it should be considered only as event which is very like to occur.

1.2.4. Functions of random variable

Let us consider an example.

Example 1.

Suppose that random varialbe X may take on two values $x_1 = 5$ and $x_2 = 10$ with the probabilities $P(X = 5) = \frac{1}{3}$ and $P(X = 10) = \frac{2}{3}$. Find the distribution function of Y where $Y = 2X$.

Solution.

Given $x_1 = 5$ & $x_2 = 10$, then the random variable Y can also take two values,

$$y_1 = 2x_1 = 10 \ \ \& \ y_2 = 2x_2 = 20$$

Where,

$$P(Y = y_1) = P(2X = 10) = P(X = 5) = \frac{1}{3},$$

$$P(Y = y_2) = P(2X = 20) = P(X = 10) = \frac{2}{3}$$

Thus Y takes on values $y_i = 2x_i (i = 1,2)$ with the same probabilities as X takes on the values x_i.The distribution function of X is,

$$F(x) = \sum_{x_i<x} P_i$$

If $x \leq 5$,

$$F(x) = \sum_{x_i<5} P_i$$

$$= \sum_{x_i<5} P(X = x_i)$$

$$= P(X = 1) + P(X = 2) + P(X = 3) + P(X = 4)$$

$$F(x) = 0 \ \ if \ x \leq 5$$

If $5 < x \leq 10$,

F$or \ x = 6$:

$$F(x) = \sum_{x_i<6} P_i$$

$$= P_1 + P_2 + P_3 + P_4 + P_5$$

$= \frac{1}{3}$

For $x = 7$:

$$F(x) = \sum_{x_i<7} P_i = P_1 + P_2 + \cdots .. + P_6 = \frac{1}{3}$$

For $x = 8$:

$$F(x) = \sum_{x_i<8} P_i = P_1 + P_2 + \cdots .... + P_7 = \frac{1}{3}$$

For $x = 9$:

$$F(x) = \sum_{x_i<9} P_i = P_1 + \cdots ... + P_8 = \frac{1}{3}$$

For $x = 10$:

$$F(x) = \sum_{x_i<10} P_i = P_1 + \cdots ... + P_9 = \frac{1}{3}$$

$\therefore F(x) = \frac{1}{3}\ \ if\ 5 < x \le 10$

If $x > 10$,

$x = 11$:

$$F(x) = \sum_{x_i<11} P_i = P_1 + P_2 + \cdots .... + P_{10} = \frac{1}{3} + \frac{2}{3} = 1$$

.

.

.

$\therefore F(x) = 1$ if $x > 10$

Hence $F(x) = \begin{cases} 0 & if\ x \le 5 \\ \frac{1}{3} & if\ 5 < x \le 10 \\ 1 & if\ x > 10 \end{cases}$

Let the distribution function of Y be $F_1(y)$. Then we have

$F_1(y) = P(Y < y)$

$= P(2X < y)$

$$= P\left(X < \frac{y}{2}\right)$$

$$i.e., F_1(y) = F\left(\frac{y}{2}\right)$$

$$\therefore F_1(y) = \begin{cases} o \ \ if \ \frac{y}{2} \leq 5 \\ \frac{1}{3} \ \ if \ 5 < \frac{y}{2} \leq 10 \\ 1 \ \ \ if \ \frac{y}{2} > 10 \end{cases}$$

$$F_1(y) = \begin{cases} 0 \ \ if \ y \leq 10 \\ \frac{1}{3} \ \ if \ 10 < y \leq 20 \\ 1 \ \ \ if \ y > 20 \end{cases}$$

Remark.

If $Y = g(X)$ is a single – valued and continuous transformation of a random variable X. Then Y is also a random variable whose distribution function is obtained from the distribution function of X.

Theorem 1.10.

Let $F(x)$ be a distribution function of a random variable X. Find the distribution function of $Y = -X$

Proof.

Let $F(x)$ be a distribution function of a random variable X.

Consider the transformation $Y = -X$

Let the distribution function of Y be $F_1(y)$. We have

$$F_1(y) = P(Y < y)$$

$$= P(-X < y)$$

$$= P(X > -y)$$

$$= 1 - P(X \leq y) \ \ldots (1)$$

If the random variable X is of the continuous type, $F(x) = \int_{-\infty}^{x} f(x)dx$

$(1) \Rightarrow F_1(y) = 1 - P(X \leq -y)$

$= 1 - F(-y)$

$\therefore F_1(y) = 1 - \int_{-\infty}^{-y} f(x)dx$

The density of Y is denoted by $f_1(y)$

Then, $F_1(y) = 1 - F(-y)$

$F_1'(y) = 0 - F'(-y)(-1)$

$F_1'(y) = F'(-y) = f(-y)$

$F_1'(y) = f(-y)$

$f_1(y) = f(-y)$

Suppose X is of the discrete type and $-y$ is its jump point,

$(1) \Rightarrow F_1(y) = 1 - P(X \leq -y)$

$= 1 - [p(X < -y) + P(X = -y)]$

$= 1 - P(X < -y) - P(X = -y)$

$F_1(y) = 1 - F(-y) - P(X = -y).$

Theorem 1.11.

Find the distribution of general linear transformation $Y = aX + b,$ where X is a random variable.

Proof.

Let $Y = aX + b$ where X is a $r.v$

Let $F_1(y)$ be distribution function of Y

Case (i): If $a > 0$, then

$F_1(y) = P(Y < y)$

$= P(aX + b < y)$

$= P(aX < y - b)$

$$= P\left(X < \frac{y-b}{a}\right)$$

$$F_1(y) = F\left(\frac{y-b}{a}\right) \quad \dots (1)$$

Suppose $f_1(y)$ is the density function,

$$f_1(y) = F_1'(y) = F'\left(\frac{y-b}{a}\right)\left(\frac{1}{a}\right)$$

$$f_1(y) = \frac{1}{a}F'\left(\frac{y-b}{a}\right)$$

$$f_1(y) = \frac{1}{a}f\left(\frac{y-b}{a}\right) \quad \dots (2)$$

$F_1'(y) = \frac{1}{a}f\left(\frac{y-b}{a}\right)$ is valid for discrete as well as continuous.

Case (ii): If $a < 0$, then

$$F_1(y) = P(Y < y)$$

$$= P(aX + b < y)$$

$$= P(aX < y - b)$$

$$= P\left(X > \frac{y-b}{a}\right) \quad (\because a < 0)$$

$$F_1(y) = 1 - P\left(X \leq \frac{y-b}{a}\right) \quad \dots (3)$$

If the random variable X is of the continuous type, then

$$F_1(y) = 1 - F\left(\frac{y-b}{a}\right)$$

Let $f_1(y)$ be the density function of Y,

$$f_1(y) = F_1'(y)$$

$$= 0 - F'\left(\frac{y-b}{a}\right)\left(\frac{1}{a}\right)$$

$$f_1(y) = \frac{-1}{a}f\left(\frac{y-b}{a}\right) \quad \dots (4)$$

From (2) & (4),

$f_1(y) = \frac{1}{|a|}f\left(\frac{y-b}{a}\right)$ if X is its type.

If the $r.v$ X is discrete type

$$(3) \Rightarrow F_1(y) = 1 - \left[P\left(X < \frac{y-b}{a}\right) - P\left(X = \frac{y-b}{a}\right)\right]$$

$$= 1 - P\left(X < \frac{y-b}{a}\right) - P\left(X = \frac{y-b}{a}\right),$$

if the point $\frac{y-b}{a}$ is a jump point of X.

At the remaining points $P\left(X = \frac{y-b}{a}\right) = 0$.

Theorem 1.13.

Find the distribution function of $Y = X^2$, swhere X is a random variable.

Proof.

Let X be a $r.v$ with distribution function $F(x)$

Given $Y = X^2$

$\Rightarrow Y$ does not take on positive value

Let $F_1(y)$ be distribution function of Y

Then,

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(Y < y) & for\ y > 0 \end{cases}$$

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(X^2 < y) = P(X < \pm\sqrt{y}) = P(-\sqrt{y} < X < \sqrt{y}) & for\ y > 0 \\ & \end{cases}$$

$$\therefore F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(-\sqrt{y} < X < \sqrt{y}) & for\ y > 0 \end{cases}$$

If the random variable X is of its type,

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ F(\sqrt{y}) - F(-\sqrt{y}) & for\ y > 0 \end{cases}$$

If the $random\ variable\ X$ has density $f(x)$,

$$f_1(y) = F_1'(y)$$

$$= \begin{cases} 0 & for\ y \leq 0 \\ F'(\sqrt{y})\frac{1}{2}\ y^{-\frac{1}{2}} + F'(\sqrt{y})\frac{-1}{2}y^{-\frac{1}{2}} & for\ y > 0 \end{cases}$$

$$f_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ \frac{F(\sqrt{y}) + F'(-\sqrt{y})}{2\sqrt{y}} & for\ y > 0 \end{cases}$$

If random variable is of discrete type,

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(X < \sqrt{y}) - P(X \leq -\sqrt{y}) & for\ \ y > 0 \end{cases}$$

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(X < \sqrt{y} - [P(X < -\sqrt{y}) + P(X = -\sqrt{y})] & for\ y > 0 \end{cases}$$

$$F_1(y) = \begin{cases} 0 & for\ y \leq 0 \\ P(X < \sqrt{y}) - P(X < -\sqrt{y}) - P(X = -\sqrt{y}) & for\ y > 0 \end{cases}$$

If the point $-\sqrt{y}$ is not a jump point of X then $P(X = -\sqrt{y}) = 0$

Exercise.

1. If X is a $r.v$ is $|X|$ is a random variable too?
2. The probability function of the $r.v$ X is of the form $P(X - r) = \frac{e^{-\lambda}\lambda^r}{r!}$ $(r = 0,1,2 \dots\dots)$. find the probability function of random variable (a) $Y = -X$ (b) $Y = aX + b$ (c) $Y = X^2$ (d) $Y = \sqrt{X}$ (e) $Y = X^l$ $(l\ is\ an\ integer)$

Remark.

Let $x_1, x_2, \dots\dots$ be the jump points of the $r.v$ X and $y_1, y_2, \dots\dots$ be the points corresponding to them according to the relation $y_i = x_i^2$

$$\therefore P(Y = y_i) = P(X^2 = y_i) = P(X = -\sqrt{y_i}) + P(X = \sqrt{y_i})$$

Example 2.

Suppose that the random variable X take an only two values$x_1 = -1$ & $x_2 = 1$, where $P(= -1) = P(x = 1) = \frac{1}{2}$. Find the distribution function of $Y = X^2$.

Solution.

Let $Y = X^2$

Since $y_1 = (-1)^2 = 1; y_2 = (1)^2 = 1$.

∴ the random variable Y taken an only one value $y = 1$.

$\therefore P(Y = 1) = P(X^2 = 1)$

$= P(X = \pm 1)$

$= P(X = +1) + P(X = -1)$

$= \frac{1}{2} + \frac{1}{2}$

$= 1$

Note: Let X be a random variable of the continuous type with density $f(x)$.

Consider a one-one transformation defined by a function $y = g(x)$ which has an everywhere continuous derivative $g(x)$.

Let (x_1, x_2) be an interval such that $g(x) \neq 0$ for $x_1 \leq x < x_2$.

Let $y_1 = g(x_1)$ & $y_2 = g(x_2)$

Let $x = h(y)$ be the function inverse to the function $g(x)$.

By our assumption, $h(y)$ is finite and continuous valued and its derivative $h'(y)$ is finite and continuous in (y_1, y_2).

$P(x_1 \leq x < x_2) = \int_{x_1}^{x_2} f(x)dx$

$$= \int_{y_1}^{y_2} f\big(h(y)\big)h'(y)dy \ldots \ldots (1)$$

Case(i): If $h'(y) > 0$, then $y_1 < y_2$.

(1)⇒ $P(x_1 \leq X < x_2) = \int_{y_1}^{y_2} f\big(h(y)\big)h'(y)dy$

$= P(y_1 \leq x < y_2)$, where $Y = g(X)$.

Case(ii): If $h'(y) < 0$, then $y_2 < y_1$.

(1)⇒ $P(x_1 \leq X < x_2) = \int_{y_1}^{y_2} f\big(h(y)\big)h'(y)dy$

$$= -\int_{y_2}^{y_1} f(h(y))h'(y)dy$$

∴ a random variable $y = g(x)$ has the density function $g(x) = f[h(y)]|h'(y)|$.

1.2.5. Multidimensional Random Variable

The following example illustrates the notion of a multidimensional random variable.

Example 1. The following table contains the data concerning the distribution of the population Poland according to sex and age from the census of 1931.

Age Group	Men	Women
0-4	2020	1962
5-9	2005	1962
10-14	1405	1372
15-19	1474	1562
20-29	2931	3213
30-39	1999	2255
40-49	1391	1596
50-59	1052	1201
60-69	753	875
70 or more	386	474
Total	15,416	16,472

The element of investigation is an inhabitant Poland in the year 1931. Every inhabitant of Poland is categorised in the table by two categories sex and age. We can assign values to this characteristic. To analyse the result of census, IBM cards are prepared for every person included in the census. To every characteristic and the consideration, a number is assigned on this card. To every man the number 1 is usually assigned and to every women the number 0. Similarly to every age group a certain number is assigned.

Consider the random event that a card chosen at random correspond to a person belonging to a given and age group.

∴ To every elementary event there correspond a pair of numbers.

Definition.

The collection of n real single-valued function $X = (X_1, X_2, \dots, X_n)$ defined on E is called $\boldsymbol{n-dimensional\ random\ variable}$ if the inverse image A of every generalized n-dimensional interval I of the form $(-\infty, -\infty, \dots, -\infty, a_1, a_2, \dots, a_n)$ is random event.

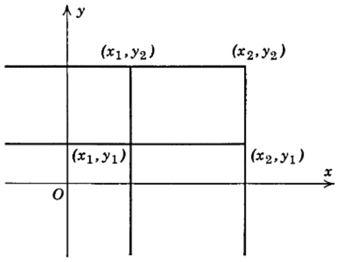
Definition.

The function $F(x, y)$ is defined by $F(x, y) = P(X < x, Y < y)$ is called the ***distribution function*** the random variable (X, Y).

Remark.

1) $P(x \le X < x_2, y_1 \le Y < y_2) = P(X < x_2, Y < y_2) - P(X < x_2, Y < y_1) - P(X < x_1, Y < y_2) + P(X < x_1, Y < y_2)$

   $= F(x_2, y_2) - F(x_2, y_1) - F(x_2, y_2) + F(x_1, y_1)$



2) The function $F(x, y)$ be the distribution function of a two-dimension random variable, it is not difficult that this function be continuous from the left non-decreasing with respect to each of the variables and satisfies the inequalities $F(\infty, \infty) = 1, F(-\infty, y) = 0, F(x, -\infty) = 0$.

   If $F(x, y)$ is a distribution function then for all values $x_2, x_1\ (x_2 > x_1)$& $y_2, y_1 (y_2 > y_1)$ the relation $F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \ge 0$ must be satisfied.

Example 2.

Let F be a function of two variables X and Y defined by $F(x, y) = \begin{cases} 0 \ if\ x + y \le 0 \\ 1 \ if\ x + y > 0 \end{cases}$. Verify that F is a distribution function or not.

Solution.

Clearly this function is non-decreasing continuous from the left with respect to x and y.

$F(-\infty, y) = F(x, -\infty) = 0, F(+\infty, +\infty) = 1.$

$$P(-1 \le X < 3, -1 \le Y < 3) = F(3,3) - F(+3, -1) - F(-1,3) + F(-1, -1)$$

$$[\because P(x_1 \le x < x_2, y_1 \le Y < y_2) = F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x, y_2)]$$

$$= 1 - 1 - 1 + 0$$

$$= -1 \ngeq 0$$

The inequality is not true for this value.

$\therefore F_1$ is not a distribution function.

Theorem 1.13.

A real-valued function $F(x, y)$ is a distribution function of a certain two dimensional random variable iff $F(x, y)$ is non-decreasing and continuous atleast from left with respect to both x and y, satisfies the inequality $F(-\infty, y) = F(x, -\infty) = 0$, $F(\infty, \infty) = 1$ and the inequality $F(x_2, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1) \geq 0$ holds for every (x_1, y_1) & (x_2, y_2), where $x_1 < x_2$ & $y_1 < y_2$.

Definition.

The two-dimensional random variable (X, Y) is said to be ***discrete type*** if, with probability 1, it takes on pairs of values belonging to a set S of pair that is at most countable and every pair (x_1, y_k) is taken with positive probability p_{ik}. These pairs of values ***jump points*** and, their probability, ***jumps***

i.e., $\sum_i \sum_k P_{ik} = 1$

The distribution function $F(x, y)$ has the form

$$F(x, y) = \sum_{\substack{x_i < x \\ y_k < y}} P_{ik}$$

where the summation is extended over all points (x_i, y_k) for which the inequalities $x_i < x$ and $y_k < y$ are satisfied.

Definition.

Let (x_i, y_k) where $i = 1,2, \dots \dots and\ k = 1,2, \dots \dots$ be an arbitrary jump point of the random variable (X, Y) of the discrete type. The probability that the random variable (X, Y) will take on the pair of values (x_i, y_k) is called the ***probability function*** of (X, Y). We write

$$P(X = x_i, Y = y_k) = P_{ik}$$

Definition.

The two-dimensional random variable (X, Y) is called as the ***continuous type,*** if there exists a non-negative function $f(x, y)$ such that for every pair (x, y) of real numbers the following relation is satisfied:

$$F(x, y) = \int_{-\infty}^{x}\left[\int_{-\infty}^{y} f(x, y)dy\right]dx$$

where $F(x, y)$ is the distribution function of (X, Y) the function $f(x, y)$ is called the ***density function***.

Definition.

The density function $F(+\infty, +\infty) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y)dx\,dy = 1$.

Definition.

If the density function $f(x, y)$ is continuous at the point (x, y), $\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x, y)$.

Remark.

If the function $f(x, y)$ is continuous of (x, y) we have,

$$f(x, y) = \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{F(x+\Delta x, y+\Delta y) - F(x,y)}{\Delta x \Delta y}$$

$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P(x<x+\Delta x, Y<y+\Delta y) - P(X<x, Y<y)}{\Delta x \Delta y}$$

$$= \lim_{\substack{\Delta x \to 0 \\ \Delta y \to 0}} \frac{P(x \le X < x+\Delta x, y \le Y < y+\Delta y)}{\Delta x \Delta y}$$

Result.

A function $F(x_1, x_2, \ldots\ldots x_n)$ is the joint distribution function of some n dimensional random variable iff F is non decreasing & continuous from the left with respect to all the arguments $x_1, x_2, \ldots\ldots x_n$ and satisfies the following conditions

i) $F(-\infty, x_2, \ldots\ldots x_n) = F(x_1, -\infty, x_3, \ldots\ldots x_n) = \cdots.. = F(x_1, x_2, \ldots\ldots -\infty) = 0$

$F(+\infty, +\infty, \dots\dots . +\infty) = 1$

ii) For all $(x_1, x_2, \dots\dots., x_n) \in R_n$ and for all $h_i > 0$ $(i = 1,2, \dots., n)$ the iuequality

$$P(x_1 \le X_1 \le x_1 + h_1, \dots\dots, x_n \le X_n < x_n + h_n) = F(x_1 + h_1, \dots\dots x_n + h_n)$$

$$\sum_{i=1}^{n} F(x_1 + h_1, \dots\dots x_{i-1} + h_{i-1}, x_i, x_{i+1} + h_{i+1}, \dots . x_n + h_n) + \sum_{\substack{i,j=1 \\ l<j}}^{n} F(x_1 + h_1, \dots\dots x_i, \dots\dots\dots + x_j, \dots . x_n + h_n) + (-1)^n F(x_1, \dots\dots x_n) \ge 0$$

Remark. Difference between thee one dimensional & multi-dimensional distribution function:

If one dimensional random variable X does not have jump points then its distribution function $F(x)$ is continuous everywhere.

But the distribution function, $F(x_1, x_2, \dots\dots x_n)$ can have discontinuity points even if the random variable $(X_1, X_2, \dots\dots . X_n)$ does not have jump points.

This is possible if $X_{j_1} = a_1, X_{j_2} = a_2, \dots\dots X_{j_r} = a_r$, where $1 \le r \le n$ and $a_1, a_2, \dots . a_n$ are constants such that $P\left(X_{j_1} = a_1, X_{j_2} = a_2, \dots\dots X_{j_r} = a_r\right) > 0$.

Example 3.

Let us consider the random vector (X, Y) with distribution function $F(x, y)$ of the form

$$F(x,y) = \begin{cases} 0 & \textit{in the domains } (-\infty < x \le 0, -\infty < y < +\infty \textit{ and } (0 < x < +\infty, -\infty < y \le 0 \\ xy & \textit{in the domain } \left(0 \le x \le 1, 0 \le y \le \frac{1}{2}\right) \\ \frac{x}{2} & \textit{in the domain } (0 \le x \le 1, \frac{1}{2} \le y < \infty \\ y & \textit{in the domain } (1 < x < \infty, 0 \le y \le 1) \\ 1 & \textit{in the domain } (x > 1, y > 1) \end{cases}$$

The random vector (X, Y) takes on with probability $\frac{1}{2}$, appoint in $\left(x = 1, \frac{1}{2} \le y \le 1\right)$

Clearly, every point with coordinates $(1, y)$ where $\frac{1}{2} \le y < \infty$ is a discontinuity point of distribution function $F(x, y)$, although the random vector (X, Y) has no jump points.

Definition.

If all the vertices of the generalized interval, given by $x_1 \leq X_1 < x_1 + h_1, x_2 \leq X_2 < x_2 + h_2, \dots \dots .., x_n \leq X_n < x_n + h_n$ are continuity points of the distribution function $F(x_1, x_2, \dots \dots x_n)$ for the surface S of the interval we have

$$P[(x_1, x_2, \dots \dots, x_n) \in S] = 0 \ \dots (1)$$

Definition. An interval, generalized or in the usual sense, for which the relation (1) is called a ***continuity interval***.

In the above example the rectangle with vertices $(1,2), (1,3), (2,2)$ and $(2,3)$ is a continuity interval by the above definition

But the vertices $(1,2)$ and $(1,3)$ are discontinuity points of distribution function $F(x, y)$.

1.2.6. Marginal distribution

Let (X, Y) be a two-dimensional random variable of the discrete type which can take on the values (x_i, y_k). Then we have

$$P(X = x_i, Y = y_k) = P_{ik}$$

Define,

$P_{.k} = \sum_i P_{ik}\, , P_{i.} = \sum_k P_{ik} \quad \dots (1)$

We have

$P_{.k} = \sum_i P_{ik}$

$\quad = P(X = x_1, Y = y_k) + P(X = x_2, Y = y_k) + \cdots$

Hence, $P_{.k} = P(Y = y_k)$ when X takes on any of the possible values.

Furthermore, it is obvious that,

$\sum_i P_{.k} = \sum_k \sum_i P_{ik} = 1$

The collection of numbers $P_{.k}$ is then a set of jumps of a probability function, the distribution determined by these jumps is called marginal distribution of random variable Y.

$\left(i.e., P_{.k} = P(Y = y_k)\right)$

The collection of numbers P_i is then a set of jumps of a probability function. The distribution determined by these jumps is called ***Marginal distribution*** of random variable Y.

$\left(i.e., P_{i.} = P(X = x_i)\right).$

Example 1.

Suppose that we have 21 slips of paper. An each slip one of the number $1,2, \dots \dots 21$ is written and there are no two slips marked with the same number. Find the marginal distribution of divisibility by 3 & 2.

Solution.

Let us assign to the appearance of a even number the number 1 and to the appearance of an odd number the number 0.

Let the random variable X takes on two values $x_1 = 1$ & $x_2 = 0$

Let the random variable Y takes on the value $y_1 = 1$ when a number divisible by 3 is chosen and the value $y_2 = 0$ otherwise

Among the 21 numbers, we have the following types of numbers:

	Number divisible by 2 x_1	Number divisible by 2 x_2	Total
No. divisible by 3	3	4	7
No. indivisible by 3	7	7	14
Total	10	11	21

Probability of divisible by 2 & 3

$$P_{11} = P(X = 1, Y = 1) = \frac{3}{21}$$

Probability of divisible by 2 not divisible by 3

$$P_{12} = P(X = 1, Y = 0) = \frac{7}{21}$$

Probability of not divisible by 2 and not divisible by 3

$$P_{22} = P(X = 0, Y = 0) = \frac{7}{21}$$

Marginal distribution of divisible by 3

$$P_{.1} = P_{11} + P_{21} = \frac{3}{21} + \frac{4}{21} = \frac{7}{21}$$

Marginal distribution of divisible by 2

$$P_{1.} = P_{11} + P_{12} = \frac{3}{21} + \frac{7}{21} = \frac{10}{21}$$

The marginal distribution of not divisible by 3

$$P_{.2} = P_{12} + P_{22} = \frac{7}{21} + \frac{7}{21} = \frac{14}{21}$$

Marginal distribution of not divisible by 2

$P_{2.} = P_{21} + P_{22} = \frac{4}{21} + \frac{7}{21} = \frac{11}{21}$.

Definition.

Let $F(x, y)$ be the distribution function of a 2-dimensional random variable (X, Y). The distribution function of the marginal distribution of X has the form

$$F(x, \infty) = P(X < x, Y < \infty)$$

If (X, Y) is a random variable of a discrete type takes a form

$$F(x, \infty) = \sum_{x_i < x} \sum_{k} P_{ik}$$

where the summation is extended over all the values of k, and those values of i for which $x_i < x$

If (X, Y) is a random variable of the continuous type,

$$F(x, \infty) = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(x, y) dy\right] dx$$

The density of the marginal distribution of the random variable X has the form

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy$$

Similarly, $F(\infty, y) = \int_{-\infty}^{y} [\int_{\infty}^{\infty} f(x, y) dx]\, dy$

The density of the marginal distribution of the random variable Y has the form

$$f_1(y) = \int_{-\infty}^{\infty} f(x, y) dx$$

Remark.

Let $F(x_1, x_2, \ldots \ldots, x_n)$ be the D.F of random vector $X_1, X_2, \ldots \ldots, X_n$ where $n > 2$. Then we can obtain $\binom{n}{k}$ k-dimensional marginal distribution for $k = 1,2, \ldots \ldots n-1$.

For example, the distribution function of random vector (X_1, X_2) has the form
$F(x_1, x_2, \infty, \ldots \ldots, \infty) = P(X_1 < x_1, X_2 < x_2, X_3 < \infty, \ldots \ldots ., X_n < \infty)$

1.2.7. Conditional distribution

Let (X, Y) be a two dimensional random vector of the discrete type, where X can take on the values $x_i (i = 1,2 \ldots \ldots)$& Y can take on the values $y_k (k = 1,2, \ldots \ldots .)$.

Let $P(X = x_i, Y = y_k) = P_{ik}$

Then the marginal distributions are

$$P(X = x_i) = P_{i.} = \sum_k P_{ik} \ \& \ P(Y = y_k) = P_{.k} = \sum_i P_{ik}$$

Let us define for every $i \ and \ k$ the probabilities

$$P(X = x_i | Y = y_k) = \frac{P_{ik}}{P_{.k}} \quad \ldots (1)$$

$$P(Y = y_k | X = x_i) = \frac{P_{ik}}{P_{i.}} \quad \ldots (2)$$

when y_k is fixed and x_i varies over all possible jump points, (1) is the probability functions of the random vector X of the discrete type under the condition $Y = y_k$

When x_i is fixed and y_k varies over all possible jump points, (2) is the probability functions of the random vector Y of the discrete type under the condition $X = x_i$.

R.H.S of (1) and (2) are non-negative numbers and bounded by 1

Since, $\sum_i P(X = x_i | Y = y_k) = \frac{\sum_i P_{ik}}{P_{.k}} = \frac{P_{.k}}{P_{.k}} = 1$

$\sum_k P(Y = y_k | X = x_i) = \frac{\sum_k P_{ik}}{P_{i.}} = \frac{P_{i.}}{P_{i.}} = 1.$

Definition. (Conditional distribution function of Y)

Consider the interval $[x, x+h)$ and the event $x < X < x+h$. Suppose that

$$P(x \leq X < x+h) > 0$$

For every value y and every interval $[x, x+h)$ we define the conditional probability

$$P(Y < y | x \leq X < x+h) = \frac{P(Y<y, x \leq X < x+h)}{P(x \leq X \leq x+h)}$$

$$= \frac{\int_x^{x+h}[\int_{-\infty}^{y} f(x,y)dy]\,dx}{\int_x^{x+h}\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx} \quad \ldots\ldots(1)$$

Suppose that the density function $f(x,y)$ is everywhere continuous and the marginal density $f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy$ is a cts function of x.

(1) $\Rightarrow$

$$\lim_{h \to 0} P(Y < y | x \leq X < x+h) = \lim_{h \to 0} \frac{\lim_{h \to 0} \int_x^{x+h}\left[\int_{-\infty}^{y} f(x,y)dy\right]dx}{\int_x^{x+h}\left[\int_{-\infty}^{\infty} f(x,y)dy\right]dx}$$

$$F(y \backslash x) = \frac{\int_{-\infty}^{y} f(x,y)dy}{f_1(x)} \quad \ldots\ldots(2)$$

For a fixed value of X equation (2) is called conditional distribution function of random variable Y.

If $g(y|x)$ is the density function of random variable Y

$$g(y|x) = \frac{f(x,y)}{f_1(x)}$$

From (2),

$$f_1(x)F(y|x) = \int_{-\infty}^{y} f(x,y)dy$$

$$\int_{-\infty}^{\infty} f_1(x)\, F(y|x)dx = \int_{-\infty}^{\infty}\left(\int_{-\infty}^{y} f(x,y)dy\right)dx$$

$$= F_2(y)$$

$$i.e., F_2(y) = \int_{-\infty}^{\infty} f_1(x)F(y|x)dx$$

Remark. Let us consider the 3-dimensional random vector (X_1, X_2, X_3) of the continuous with the density $f(x_1, x_2, x_3)$ which is everywhere continuous and with all the marginal densities continuous.

$$F(x_3|x_1, x_2) = \frac{\int_{-\infty}^{x_3} f(x_1, x_2, x_3)dx_3}{\int_{-\infty}^{\infty} f(x_1, x_2, x_3)dx_3}$$

$$F(x_3, x_2|x_1) = \frac{\int_{-\infty}^{x_2}\left(\int_{-\infty}^{x_3} f(x_1, x_2, x_3)dx_3\right)dx_2}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2, x_3)dx_3\, dx_2}$$

Example 1.

Suppose that we have 21 slips of papers. An each, slip one of the numbers 1,2, … … 21 is written, and there are no two slips marked with the same number. What is the probability that a no. chosen at random will be divisible by 3 given that it is even.

Solution.

Let us assign to the appearance of a even number the number 1 and to the appearance of an odd number the number 0.

Let the random variable X takes on two values $x_1 = 1$ & $x_2 = 0$

Let the random variable Y takes on the values, when a number divisible by 3 is chosen and the value $y_2 = 0$ otherwise.

Among the 21 numbers, we have

	Number divisible by 2 x_1	Number divisible by 2 x_2	Total
No. divisible by 3 y_1	3	4	7
No. indivisible by 3 y_2	7	7	14
Total	10	11	21

Probability of a no. chosen at random is divisible by 3 given that it is even.

$$P(Y=1|X=1)=\frac{P(Y=1,X=1)}{P(X=1)}$$

$$=\frac{P_{11}}{P_1}=\frac{3}{10}$$

Probability of a number chosen at random is not divisible by 3 given that it is even.

$$P(Y=0|X=1)=\frac{P(Y=0,X=1)}{P(X=1)}=\frac{7}{10}$$

Conditional distribution of odd number into those divisible and not divisible by 3.

$$P(Y=1|X=0)=\frac{P(Y=1,X=0)}{P(X=0)}=\frac{4}{11}$$

$$P(Y=0|X=0)=\frac{P(Y=0|X=0)}{P(X=0)}=\frac{7}{11}.$$

Definition.

Let X be a random variable and S is a Borel set on the x −axis such that $0<P(X\in S)<1$. The conditional distribution defined for any real Z by the expression $P(X<x|X\in S)$ is called the ***truncated distribution*** of X.

If X is of the discrete type with jump points x_i and jumps P_i, the probability function of the truncated distribution of X is of the form

$$P(X=x_i|X\in S)=\frac{P(X=x_i,X\in S)}{P(X\in S)}$$

$$S=\begin{cases}\frac{P_i}{\sum_{x_i\in S}P_j} & if\ x_i\in S\\ 0 & if\ x_i\notin S\end{cases}$$

If X is of continuous type with the density $f(x)$, then

$$P(X<x|X\in S)=\frac{P(X<x,X\in S)}{P(X\in S)}$$

$$=\frac{\int_{(-\infty,x)\cap S}f(x)dx}{\int_S f(x)dx}$$

The density $g(x)$ of this distribution takes the form

$$g(x) = \begin{cases} \dfrac{f(x)}{\int_S f(x)dx} & if\ x \in S \\ 0 & if\ x \notin S \end{cases}$$

Example 2.

Consider the random variable X with density

$$f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & x > 0, x > 1 \end{cases}$$

Let $S = \left[0, \frac{1}{2}\right)$, then we have

$$P(X \in S) = \int_0^{\frac{1}{2}} dx = \frac{1}{2}$$

The density $g(x)$ of the truncated distribution of X is of the form

$$g(x) = \begin{cases} 2 & 0 \leq x \leq \dfrac{1}{2} \\ 0 & x < 0, x > \dfrac{1}{2} \end{cases}$$

1.2.8. Independent random variables

Let $F(x, y), F_1(x)\ and\ F_2(y)$ denote respectively the two-dimensional distribution function of the random variable (X, Y) and the marginal distribution function of the random variables X and Y.

Definition.

The random variable X and Y are said to be ***independent*** proof for every pair (x, y) of real numbers the equality

$$F(x, y) = F_1(x)F_2(y) \ldots \ldots (1)$$

is satisfied.

Remark.

Let (a, b) and (c, d), where $a < c$ and $b < d$ be two arbitrary point in the plane (x, y). We have

$$P(a \leq X \leq c) = F_1(c) - F_1(a)$$

$$P(b \leq Y < d) = F_2(d) - F_2(b)$$

Multiplying the RHS and LHS of these relation and applying (1) gives,

$$P(a \leq X < c)P(b \leq Y < d)$$

$$= [F_1(c) - F_1(a)][F_2(d) - F_2(b)]$$

$$= F_1(c)F_2(d) - F_1(c)F_2(b) - F_1(a)F_2(d) + F_1(a)F_2(b)$$

$$= F(c, d) - F(c, b) - F(a, d) + F(a, b)$$

We know that,

$$P(x_1 \leq X < x_2, y_1 \leq Y < y_2)$$
$$= F(x_1, y_2) - F(x_2, y_1) - F(x_1, y_2) + F(x_1, y_1)$$

We obtain,

$$P(a \leq X < c)P(b \leq X < d) = P(a \leq X < c, b \leq Y < d) \dots\dots (2)$$

Remark.

Let us consider a tow dimensional random variable (X, Y) of discrete type with jump points (x_i, y_k) and jumps P_{ik}. Suppose that X and Y are independent.

Then in the special case of (2) for which the rectangles $(a \leq X < c, b \leq Y < d)$ and reduced to the points $(X = a, Y = b)$, we obtain the equality

$$P_{ik} = P(X = x_i, Y = y_k)$$

$$= P(X = x_i)P(Y = y_k)$$

$$= P_{i.}P_{.k} \quad \dots \dots (3)$$

for every pair (x_i, y_k). The equality (3) is satisfied if for every pair (x_i, y_k), (1) is also satisfied.

We have to shown that *if (X, Y) is a random variable of the discrete type, equality (3) holding for every pair (x_i, y_k) is a necessary and sufficient condition for the independence of the random variable X and Y.*

For every pair of numbers (i, k), we obtain

$$P(X = x_i | Y = y_k) = \frac{P(X=x_i, Y=y_k)}{P(Y=y_k)}$$

$$= \frac{P_{i.}P_{.k}}{P_{.k}} = P_{i.}$$

$$= P(X = x_i) \quad \dots \dots (4a)$$

$$P(Y = y_k, X = x_i) = \frac{P(Y=y_k, X=x_i)}{P(X=x_i)}$$

$$= \frac{P_{.k}P_{i.}}{P_{i.}} = P_{.k} = P(Y = y_k) \quad \dots \dots (4b)$$

From (4a) and (4b) it follows that, if the random variable X and Y are independent the distribution of X is same for all values of the random variable Y

Thus no value obtained for the variable Y gives any information about the distribution of the variable X and conversely, the condition distribution of the random variable Y is identically for all values of X.

Theorem 1.14.

If (X, Y) is a random variable of the continuous type with density function $f(x, y)$ is every where continuous, the validity of

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) = F'(x)F'(y) = F_1(x)F_2(y)$$

for arbitrary point (x, y) is a necessary and sufficient condition for the independent of random variable X and Y.

Proof.

If the random variable (X, Y) is of the continuous type differentiation of expression

$$F(x,y) = F_1(x)F_2(y)$$

with respect to x, y, with the possible exception of the set of points at which the density function $f(x,y)$ is not continuous, gives

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) = F_1'(x)F_1'(y) = f_1(x)f_2(y)$$

Conversely, let $\frac{\partial^2 F(x,y)}{\partial x \partial y} = f_1(x)f_2(y)$.

Then $F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(x,y)dydx$

$$= \int_{-\infty}^{x} [\int_{-\infty}^{y} f_1(x)f_2(y)dy]\, dx$$

$$= \int_{-\infty}^{x} f_1(x)dx \int_{-\infty}^{y} f_2(y)dy$$

$i.e., F(x,y) = F_1(x)F_2(y)$

$\therefore$ If (X, Y) is a random variable of the continuous type whose density function $f(x,y)$ is everywhere continuous.

Example 1.

Consider the two consecutive tosses of a coin. The random variable X takes on the value 0 or 1 according to whether heads or tails appear as a result of the first toss. The random variable Y takes on the value 0 or 1 according to whether heads or tails appear as a result of the second toss.

$\therefore$ The two-dimensional random variable (X, Y) may take on the values $(1,1), (1,0), (0,1), (0,0)$.

The probability of each of these events is the same and hence equals $\frac{1}{4}$.

sBoth $X$and Y take on the values 0 and 1 with probability $\frac{1}{2}$. Thus we have,

$$P(X = 1, Y = 1) = \frac{1}{4} = \frac{1}{2}.\frac{1}{2} = P(X = 1)P(Y = 1)$$

$$P(X = 1, Y = 0) = \frac{1}{4} = P(X = 1)P(Y = 0)$$

$$P(X = 0, Y = 1) = \frac{1}{4} = P(X = 0)P(Y = 1)$$

$$P(X = 0, Y = 0) = \frac{1}{4} = P(X = 0)P(Y = 0)$$

$\therefore X$ and Y are independent

Remark.

Let X_1 and X_2 be two independent random variables. Consider two single-valued functions $Y_1 = g_1(x_1)$ and $Y_2 = g_2(x_2)$. Y_1 and Y_2 are also random variable. Show that Y_1 and Y_2 are independent.

Proof

Let $h_1(-\infty, y_1)$ and $h_2(-\infty, y_2)$ be the Borel sets into which the functions inverse to g_1 and g_2 map the intervals $(-\infty, y_1)$ and $(-\infty, y_2)$.

$$\therefore P(Y_1 < y_1, Y_2 < y_2) = P[g_1(X_1) < y_1, g_2(X_2) < y_2]$$

$$= P[X_1 \in h_1(-\infty, y_1), X_2 \in h_2(-\infty, y_2)]$$

$$= P[X_1 \in h_1(-\infty, y_1)]P[X_2 \in h_2(-\infty, y_2)]$$

$$= P(Y_1 < y_1)P(Y_2 < y_2).$$

Definition

The random variable $X_1, X_2, \ldots\ldots, X_n$ are called ***independent*** it for n arbitrary real numbers $(x_1, x_2, \ldots\ldots, x_n)$ the following relation is satisfied:

$$F(x_1, x_2, \ldots\ldots, x_n) = F_1(x_1)F_2(x_2) \ldots\ldots F_n(x_n),$$

where $F(x_1, x_2, \ldots\ldots, x_n)$ is the distribution function of the random variable $(X_1, X_2, \ldots\ldots, X_n)$ and $F_1(x_1), \ldots\ldots F_n(x_n)$ are the marginal distribution function of $X_1, X_2, \ldots\ldots, X_n$.

Example.

If the random variable $X_1, X_2, \ldots\ldots, X_n$ are independent then for every $s \leq n$ the random variable $X_{k_1}, X_{k_2}, \ldots\ldots, X_{k_s}$, where $1 \leq k_1 < k_2 < \cdots .. < k_s \leq n$

For simplicity is notation assume that $k_1 = 1, k_2 = 2, \ldots\ldots k_s = s$

$F(x_1, x_2, \ldots\ldots, x_s + \infty_1, \ldots\ldots. +\infty)$

$= \lim_{x_{s+1}\to\infty_1, \ldots\ldots x_n\to\infty} F(x_1, x_2, \ldots\ldots, x_s, x_{s+1})$

$= \lim_{x_{s+1}\to\infty_1, \ldots\ldots x_n\to\infty} [F_1(x_1)\, F_2(x_2) \ldots\ldots F_s(x_s)F_{s+1}(x_{s+1}) \ldots\ldots F_n(x_n)]$

$= F_1(x_1)F_2(x_2) \ldots\ldots F_s(x_s)F_{s+1}(+\infty)\, F_n(+\infty)$

$= F_1(x_1)F_2(x_2) \ldots\ldots F_s(x_s).$

We now give the definition of independent of a countable no. of random variables.

Definition.

The random variable $X_1, X_2, \ldots\ldots, X_n$ are called ***independent*** if for ever $n = 2,3 \ldots\ldots$ the random variable $X_1, X_2, \ldots\ldots, X_n$ are independent.

Definition.

The random vectors $X = (X_1, X_2, \dots\dots, X_j)$ and $Y = (Y_1, Y_2, \dots\dots, Y_r)$ are ***independent*** if for every $j + r$ real numbers. $x_1, x_2, \dots\dots, x_j, y_1, y_2, \dots\dots, y_r$ we have

$$F(x_1, x_2, \dots\dots, x_j, y_1, y_2, \dots\dots, y_r) = G(x_1, x_2, \dots\dots, x_j)H(y_1, y_2, \dots\dots, y_r)$$

where $F, G \; and \; H$ are the distribution function of the random vectors $(X_1, X_2, \dots\dots, X_j, Y_1, Y_2, \dots\dots, Y_r), (X_1, X_2, \dots\dots, X_j)$ and $(Y, Y_2, \dots\dots, Y_r)$ respectively.

1.2.9. Functions of Multidimensional Random Variables

Here we give the formula for the two-dimensional probability density of a function of a random variable (X, Y) of the continuous type. Let

$$U_1 = g_1(X, Y), \quad U_2 = g_2(X, Y) \; \dots..(1)$$

be a continuous one-to-one mapping of the random variable (X, Y) with density $f(x, y)$. Suppose that the functions g_1 and g_2 have continuous partial derivatives with respect to x and y, and let $(a \leq X < b, c \leq Y < d)$ be a rectangle on which the Jacobian of the transformation (1) is different from zero.

Denote by $x = h_1(u_1, u_2)$ and $y = h_2(u_1, u_2)$ the inverse transformation. By our assumptions, the functions h_1 and h_2 are also one-to-one and have continuous partial derivatives with respect to u_1 and u_2. Denote by J the Jacobian of the inverse transformation

$$J = \begin{vmatrix} \frac{\partial x}{\partial u_1} & \frac{\partial x}{\partial u_2} \\ \frac{\partial y}{\partial u_1} & \frac{\partial y}{\partial u_2} \end{vmatrix}$$

By our assumptions, this Jacobian is finite and continuous in the domain S of the plane (u_1, u_2), where S is the image of the rectangle $(a \leq \; X \; < \; b, c \leq \; Y \; < \; c)$ given by transformation (l). We have

$$P(a \leq X < b, c \leq Y < d) = \int_a^b \left[\int_c^d f(x, y) dy \right] dx$$

$$= \iint_{(S)} f[h_1(u_1, u_2), h_2(u_1, u_2)]|J|du_1du_2 \ldots..(2).$$

It follows from (2) that the two-dimensional density of the random variable (U_1, U_2) has the form

$$r(u_1, u_2) = f[h_1(u_1, u_2), h_2(u_1, u_2)]|J| \ldots \ldots (3)$$

We now investigate the distribution of the sum, difference, product, and ratio of two random variables. They are given here as examples of continuous functions of multidimensional random variables, but at the same time the formulas involved are very important in probability theory and its applications.

Example 1.

Consider again two consecutive throws of a die. Let the random variable X correspond to the result of the first throw and Y to the result of the second throw. The random variables X and Y are independent. Both X and Y take on the values $1, \ldots \ldots, 6$ each with probability $\frac{1}{6}$. Hence the two dimensional random variable (X, Y) can take on the pairs of values (i, k), where i and k run over all integers from 1 to 6. Let us form the value of the sum $i + k$ for every possible pair (i, k). All possible values of the sum $i + k$ form the set of possible values of a new random variable which will be called the sum of the random variables X and Y. This sum is again a one-dimensional random variable and takes on the following values: $2,3,4, \ldots \ldots, 11,12$.

Let $Z = X + Y$. We shall compute the probability function of Z. Because of the independence of X and Y we have

$P(Z = 2) = P(X = 1)P(Y = 1) = \frac{1}{36}$,

$P(Z = 3) = P(X = 1, Y = 2) + P(X = 2, Y = 1) = \frac{1}{18}$,

$\mathrm{P}(\mathrm{Z} = 4) = P(X = 1, Y = 3) + P(X = 2, Y = 2) + P(X = 3, Y = 1) = \frac{1}{12}$,

$\ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots$

$\mathrm{P}(\mathrm{Z} = 12) = P(X = 6, Y = 6) = \frac{1}{36}$.

Then

$$\mathrm{P}(\mathrm{Z} = 2) + P(Z = 3) + \cdots \ldots . + P(Z = 12) = 1$$

In the example of the double throw of a die we could also have considered the random variable $\mathrm{V} = \mathrm{X} - \mathrm{Y}$. The set of all possible values of this random variable consists of all possible values of the difference of the numbers i and k. The random variable V then takes on the values

$$-5, -4, \dots\dots, 0, 1, 2, \dots\dots., 5.$$

For example

$$P(V = 0) = P(X = 1, Y = 1) + P(X = 2, Y = 2) + P(X = 3, Y = 3)$$

$$+P(X = 4, Y = 4) + P(X = 5, Y = 5) + P(X = 6, Y = 6) = \frac{1}{6},$$

$P(V = -5) = P(X = 1, Y = 6) = \frac{1}{36}.$

Then

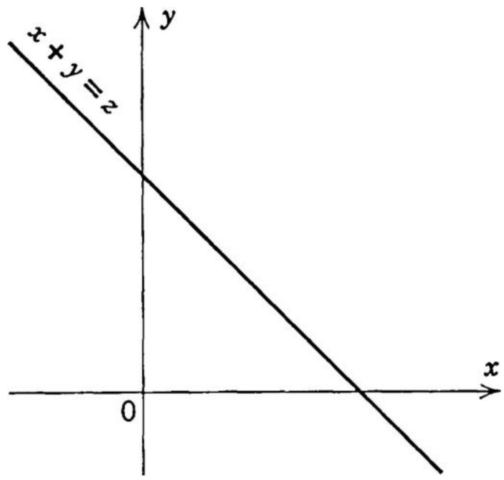
$$\Sigma_{k=-5}^{5} P(V = k) = 1$$

Similarly, we could have considered the probability function of the random variable $T = XY$ and of $S = Y|X$

We can see from this example what is understood by the sum, difference, product, and ratio of two random variables. Thus, the random variable $X + Y$ is a function of the two-dimensional random variable (X, Y). The set of possible values of $X + Y$ is formed from all possible values of the sum $x + y$, where x is a possible value of X and y is a possible value of Y. Similarly, the sum of any finite number of random variables can be defined.

The set of possible values of the random variable $X - Y$ consists of all values $x - y$, where x is a possible value of X and y is a possible value of Y. The product and ratio of two random variables are defined in an analogous way.

C Suppose we are given the distribution of the two-dimensional random variable (X, Y). We shall find the distributions of the random variables obtained as a result of four arithmetic operations performed on X and Y.



We shall find the distribution function of the sum

$$Z = X + Y \ldots..(4)$$

of two random variables or, in other words, the function

$$F(z) = P(X + Y < z).$$

Fig. 2.9.1

If for a given value of z we draw on the plane (x, y) the line x + y = z (see Fig. 2.9. l), then F(z) will be the probability that the point with coordinates x, y lies below this line.

If (X, Y) is a random variable of the discrete type and takes on values (x_i, y_k)

$$F(z) = \Sigma_{x_i+v_k<z} P(X = x_i, Y = y_k) \ldots.(5),$$

where the summation is extended over all the values (x_i, y_k) for which the inequality under the summation sign is satisfied.

Let (X, Y) be a random variable of the continuous type and let $f(x, y), f_1(x), f_2(y)$ denote respectively the densities of the random variables

(X, Y), X and Y. Let us write equality (4) in the form

$X = X, Z = X + Y$ or $X = X, Y = Z — X$,….(6)

where the identity X = X is added in order to reduce this problem to a special case of transformation (l) considered at the beginning of this section. We have

$$J = \begin{vmatrix} 1 & 0 \\ -1 & 1 \end{vmatrix} = 1$$

From (3) it follows that the density of (X, Z) is

$$f(x, z - x) \ldots.(7)$$

The density ψ(z) of Z is obtained as a marginal density of the two-dimensional random variable (X, Z) by integrating (2.9.7) with respect x to from $-\infty$ to $+\infty$,

$$\psi(z) = \int_{-\infty}^{+\infty} f(x, z - x)dx \ldots.(8)$$

Finally we obtain

$$F(z) = \int_{-\infty}^{z} \left[\int_{-\infty}^{+\infty} f(x, z - x)dx\right]dz \ldots..(9)$$

If the random variables X and Y are independent, according to (5), we have

$$f(x, y) = f_1(x) f_2(y)$$

Hence

$$\psi(z) = \int_{-\infty}^{+\infty} f_1(x) f_2(z-x) dx \quad .....(8')$$

and

$$F(z) = \int_{-\infty}^{z} \left[\int_{-\infty}^{+\infty} f_1(x) f_2(z-x) dx\right] dz \quad ....(9')$$

Because of the symmetry of the sum, we can replace in formulas (8), (8'), (9), and (9'), x by z − y and z − x by y.

Example 2. Consider the random variable X with the density

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad ....(10)$$

Instead of the expression $e^{-\frac{x^2}{2}}$ we often write $\exp\left(-\frac{x^2}{2}\right)$. Then

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = 1$$

Expression (10) is the density of the Gauss distribution, which is also called the normal distribution. This distribution will be treated more extensively later.

Let the random variable Y have the density

$$f(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

Suppose that X and Y are independent. Thus the density of the joint random variable (X, Y) is

$$f(x, y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} = \frac{1}{2\pi} e^{-\frac{(x^2+y_2)}{2}} \quad .....(11)$$

Consider the random variable Z = X + Y. By (8') we have for the density of Z

$$\psi(z) = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(z-x)^2}{2}\right) dx$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp\left(-\frac{2x^2-2zx+z^2}{2}\right) dx$$

since

$$2x^2 - 2zx + z^2 = \left(x\sqrt{2}\right)^2 - 2\left(x\sqrt{2}\right)\frac{z}{\sqrt{2}} + \left(\frac{z}{\sqrt{2}}\right)^2 + \frac{z^2}{2} = \left(x\sqrt{2} - \frac{z}{\sqrt{2}}\right)^2 + \frac{z^2}{2}$$

we have

$$-\frac{1}{2}(2x^2 - 2zx + z^2) = -\frac{1}{2}\left(x\sqrt{2} - \frac{z}{\sqrt{2}}\right)^2 - \frac{z^2}{4},$$

and hence

$$\psi(z) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \exp[-\frac{1}{2}\left(x\sqrt{2} - \frac{z}{\sqrt{2}}\right)^2 \exp\left(-\frac{z^2}{4}\right)]\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{4}\right) \frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} exp\left[-\frac{1}{2}\left(x\sqrt{2} - \frac{z}{\sqrt{2}}\right)^2\right] dx$$

Introducing the substitution $u = x\sqrt{2} - z|\sqrt{2}$ into the last integral we obtain

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{+\infty} exp\left[-\frac{1}{2}\left(x\sqrt{2} - \frac{z}{\sqrt{2}}\right)^2\right] dx = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} \exp\left(-\frac{u^2}{2}\right) du = 1$$

and finally

$$\psi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{4}\right)$$

Thus the distribution function F(z) is given by the formula

$$F(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{z^2}{4}}\, dz$$

We leave it to the reader as an exercise to derive the formulas for the distribution function and the density of the difference of two random variables.

Let us now consider the product of two random variables X and Y with the two-dimensional density $f(x, y)$ and the marginal densities $f_1(x)$ and $f_2(y)$ respectively. Let

$$Z = XY \ldots..(12)$$

This equality may be written as a system of equalities

$$X = X, \quad Y = \frac{Z}{X}$$

We have

$$J = \begin{vmatrix} 0 & 1 \\ \frac{1}{x} & -\frac{z}{x^2} \end{vmatrix}$$

and hence $|J| = 1|x|$. It follows from (3) that the two-dimensional density of the random variable (Z, X) has the form

$$f\left(x, \frac{z}{x}\right)\frac{1}{|x|} \ldots.(13)$$

The density $\psi(z)$ of Z can be obtained by integrating expression (2.9.13) from $-\infty$ to $+\infty$. We have

$$\psi(z) = \int_{-\infty}^{+\infty} f\left(x, \frac{z}{x}\right)\frac{1}{|x|}dx \ldots.(14)$$

The distribution function of Z has the form

$$F(z) = \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} f\left(x, \frac{z}{x}\right)\frac{1}{|x|}\right]dz \ldots..(15)$$

If the random variables X and Y are independent, equalities (14) and (15) can be written as

$$\psi(z) = \int_{-\infty}^{+\infty} \frac{1}{|x|} f_1(x) f_2\left(\frac{z}{x}\right)dx \ldots.(14')$$

and

$$F(z) = \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} \frac{1}{|x|} f_1(x) f_2\left(\frac{z}{x}\right)dx\right]dz \ldots.(15')$$

we can replace in (14), (14'), (5), and (15'), x by z/y, z/x by y, and |x| by |y|

For the ratio of two random variables,

$$Z = \frac{X}{Y},$$

we obtain the following formulas:

$$\psi(z) = \int_{-\infty}^{+\infty} f(yz, y)|y|dy \ldots..(16)$$

and

$$F(z) = \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} f(yz, y)|y|dy\right]dz \ldots.(17)$$

If the random variables X and Y are independent,

$$\psi(z) = \int_{-\infty}^{+\infty} f_1(yz)\, f_2(y)|y|dy$$

$$F(z) = \int_{-\infty}^{z}\left[\int_{-\infty}^{+\infty} f_1(yz) f_2(y)\ |y|dy\right]dz.$$

UNIT II

PARAMETERS OF THE DISTRIBUTION OF A RANDOM VARIABLE

2.1. Expected values

With every distribution of a random variable there are associated certain numbers called the ***parameters of the distribution***, which play an important role in mathematical statistics. The parameters of a distribution are the moments and functions of them and also the order parameters.

Let X be a random variable. Consider a single-valued function $g(X)$ of X.

Definition.

Let X be a random variable of a discrete type with jump points x_k and jumps p_k. The series

$$E[g(X)] = \sum_k p_k g(x_k) \quad \dots\dots (1)$$

is called the expected value of the random variable $g(X)$ if the following inequality is satisfied:

$$\sum_k p_k |g(x_k)| < \infty \qquad \dots\dots (2)$$

Definition.

Let X be a random variable of the continuous type with density function $f(x)$. Let $g(x)$ be the Riemann integral. The integral

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx \dots\dots (3)$$

is called the ***expected value*** of the random variable $g(X)$ if the following inequality is satisfied:

$$\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty \dots\dots (4)$$

Remark.

1. When the R.H.S of (1) is exists but (2) is not satisfied then the expected value of the random variable $g(X)$ does not exists.
2. Similarly, when the R.H.S of 3 is exists but (4) is not satisfied then the expected value of the random variable $g(X)$ does not exists.
3. If $Y = g(X)$, the expected value $E[g(x)]$ equals the expected value $E(Y)$ computed directly from the distribution of the random variable Y.

Suppose X is a random variable of a discrete type. Let X have jump points x_k and jumps p_k and let Y have jump points y_j and jumps q_j.

Notice that $q_j(j = 1,2, \dots.)$ Equals the sum of the probabilities p_k for those k for which the equality $g(x_k) = y_i$ holds.

Since by the assumption about the existence of expectation. $E[g(X)]$, the series $\sum_R p_k g(x_k)$ is absolutely convergent.

$\therefore E(Y) = \sum_j q_j y_j = \sum_K p_k g(x_k) = E[g(X)]$.

Example 1.

Suppose that the random variable X can take on two values $x_1 = -1$ with probability $P_1 = 0.1$ and $x_2 = 1$ with probability $p_2 = 0.9$. find the $E(x)$

Solution.

$$E(X) = \sum_K p_k x_k$$

$$= \sum_{K=1}^{2} p_k x_k$$

$$= p_1 x_1 + p_2 x_2$$

$$= (0.1)(-1) + (0.9)(1)$$

$$= -0.1 + 0.9$$

$$= 0.8$$

Clearly, $\sum_{k=1}^{n} p_k |g(x_k)| < \infty$ $[since\ \sum_{k=1}^{n} p_k |g(x_k)| = P_1|x_1| + P_2\,|x_2| = 0.1 \times 1 + 0 = 1 < \infty]$.

Example 2.

Find the expectation value of Poisson Distribution (or) Let the random variable X takes on the value $K = 0,1,2, \dots \dots$ and Let $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$ where $\lambda > 0$ is a positive constant. Compute expected value of X.

Solution.

$$E(X) = \sum_{k=0}^{\infty} k\ \frac{\lambda^k}{k!} e^{-\lambda}$$

$$= 0 + \lambda e^{-\lambda} \sum_{k=1}^{\infty}\ \frac{k\lambda^{k-1}}{k(k-1)!}$$

$$= \lambda e^{-\lambda} \sum_{r=0}^{\infty}\ \frac{\lambda^r}{r!} \quad (put\ r = k - 1)$$

$$= \lambda e^{-\lambda} \left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \cdots \dots .\right)$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$E(X) = \lambda$$

Clearly, $\sum_{K=0}^{\infty} p_k |K| < \infty$.

Example 3.

The random variable X takes on the values $r = 0,1,2, \dots \dots n$ with $P(X = r) = \frac{n!}{r!(n-r)!} P^r (1 - P)^{n-r}$. Find $E(X)$.

Solution.

$$E(X) = \sum_K x_k p_k$$

$$E(X) = \sum_{r=0}^{n} r p_r$$

$$= \sum_{r=0}^{n} r \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

$$= 0 + \sum_{r=1}^{n} r \frac{n(n-1)!}{r(r-1)!(n-r)!} p p^{r-1} (1-p)^{n-r}$$

$$= np \sum_{r=1}^{n} \frac{(n-1)!}{(r-1)!(n-r)!} p^{r-1} (1-p)^{n-r}$$

$$= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!\big(n-(k+1)\big)!} P^k (1-p)^{n-(k+1)}$$

$$= np\, [1 - p + p]^{n-1} = np[1]^{n-1} = np$$

Clearly $\sum_{r=0}^{n} p_k |r| < \infty$.

Example 4.

sA random variable X is of the continuous type with density function $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Find $E(X)$.

Solution.

$$E(X) = \int_{-\infty}^{\infty} x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x \; e^{-\frac{x^2}{2}} dx$$

$$= \frac{-1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} (-xdx)$$

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

$$= \frac{-1}{\sqrt{2\pi}} \left(e^{-\frac{x^2}{2}}\right)_{-\infty}^{\infty} = \frac{-1}{\sqrt{2\pi}} (e^{-\infty} - e^{-\infty})$$

$E(X) = 0.$

Example 5.

Let the random variable X take on the values $x_k = \frac{(-1)^k 2^k}{k}, (k = 1,2,3 \ldots\ldots), p_k = \frac{1}{2^k}$. Find $E(X)$

Solution.

$$E(X) = \sum_k p_k x_k$$

$$= \sum_{k=1}^{\infty} \frac{1}{2^K} \frac{(-1)^K 2^K}{K} = \sum_{k=1}^{\infty} \frac{(-1)^K}{K}$$

$$= \frac{(-1)^1}{1} + \frac{(-1)^2}{2} + \frac{(-1)^3}{3} + \cdots \ldots.$$

$$E(X) = -\log 2$$

But $\sum_{k=1}^{\infty} p_k |(x_k)| < \sum_{k=1}^{\infty} \frac{1}{2^K} \left|\frac{(-1)^K 2^K}{K}\right|$

$$= \sum_{k=1}^{\infty} \frac{1}{K}$$

$$= \infty$$

$\therefore E(X)$ does not exists.

Example 6.

Consider the random variable Y defined by $Y = |X|$ where X has distribution is given in above example compute $E(Y)$.

Solution.

Given $Y = |X|$

$$E(Y) = \sum_K P_k\, Y$$

$$= \sum_K P_k\, |X| = \sum_K \frac{1}{2^K}\left|\frac{(-1)^K 2^K}{K}\right|$$

$$= \sum_K \frac{1}{K} = \infty$$

$\therefore\ E(Y)$ does not exist.

2.2. Moments

Definition.

The expected value of the function $g(X) = X^K$

i.e., $m_K = E(X^K)$

is called the ***moment of order k*** of the random variable X.

Definition.

If the random variable X is of the discrete type with jump points x_l and jumps p_l.

The moment of order k is $m_K = E(X^K) = \sum_l x_l^K P_l$

Definition.

If the random variable X is of the continuous type with density function $f(x)$, the moment of order k is given by

$$m_k = E(X^k) = \int_{-\infty}^{\infty} x^k f(x)dx.$$

Note.

1. The moment m_K to exists, it is necessary that $\sum_l x_l^k P_l$ (or) $\int_{-\infty}^{\infty} x^k f(x)dx$ be absolutely convergent.
2. If the moment of order k exists all the moment of order smaller than k also exists.
3. If the moment of order k of random variable X exists then $\lim_{a\to\infty} a^k P(|X| > a) = 0 \ where \ a > 0$

Proof.

If X is a random variable of continuous type with density $f(x)$,

$$\lim_{a\to\infty} a^k P(|X| > a) \leq \lim_{a\to\infty} \int_{|x|>a} |x|^k f(x)dx = 0$$

i.e., $P(|X| > a) = 0\left(\frac{1}{a^{ks}}\right)$.

Theorem 2.1.

Let $g_1(X)$ and $g_2(X)$ be two single-valued functions of a random variable X and Let the expected values $E[g_1(X)]$ and $E[g_2(X)]$ exists. Show that $E[(X) + g_2(X)] = E[g_1(X)] + E[g_2(X)]$

Proof.

Let X be the continuous type.

Since $E[g_1(X)]$ and $E[g_2(X)]$ exists,

$\Rightarrow \int_{-\infty}^{\infty}|g_1(X)|\, f(x)dx < \infty$ and $\int_{-\infty}^{\infty}|g_2(X)|\, f(x)dx < \infty$

$\Rightarrow \int_{-\infty}^{\infty}|g_1(x)|\, f(x)dx + \int_{-\infty}^{\infty}|g_2(x)|\, f(x)dx < \infty$

$$\Rightarrow \int_{-\infty}^{\infty} [|g_1(x)| + |g_2(x)|]\, f(x)dx < \infty \quad \ldots\ldots (1)$$

We know that $|g_1(x) + g_2(x)| \leq |g_1(x)| + |g_2(x)|$,

$$\int_{-\infty}^{\infty} |g_1(x) + g_2(x)|\, f(x)dx \leq \int_{-\infty}^{\infty} [|g_1(x) + g_2(x)|]\, f(x)dx < \infty$$

$\therefore \int_{-\infty}^{\infty} |g_1(x) + g_2(x)|\, f(x)dx < \infty$

$\therefore \int_{-\infty}^{\infty} |g_1(x) + g_2(x)|\, f(x)dx < \infty$

$\therefore E[g_1(X) + g_2(X)]$ is exists.

Consider $E[g_1(X) + g_2(X)] = \int_{-\infty}^{\infty} \big(g_1(x) + g_2(x)\big)\, f(x)dx$

$$= \int_{-\infty}^{\infty} g_1(x)\, f(x)dx + \int_{-\infty}^{\infty} g_2(x)\, f(x)dx$$

$$= E\big(g_1(X)\big) + E\big(g_2(X)\big)$$

Note.

Suppose $g_1(X), g_2(X), \ldots\ldots g_n(X)$ are single valued functions of random variable X and $E[g_1(X)], E[g_2(X)], \ldots\ldots.. E[g_n(X)]$ are exists. Then

$$E[g_1(X) + g_2(X) + \cdots g_n(X)] = E[g_1(X)] + \cdots + E[g_n(X)]$$

Theorem 2.2.

Show that $E[(aX)^k] = as^k E[X^k]$ where a is constant.

Proof.

Let X be the random variable of continuous type.

Suppose, $E[X^k]$ is exists.

Consider, $E[(aX)^k] = \int_{-\infty}^{\infty} (aX)^k\, f(x)dx$

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

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

$$= a^k E(X^k)$$

$\therefore E((aX)^k) = a^k E(X^k)$

To prove: $\int_{-\infty}^{\infty}(aX)^k f(x)dx < \infty$

$\int_{-\infty}^{\infty}(aX)^k f(x)dx = a^k \int_{-\infty}^{\infty} X^k f(x)dx < \infty$

$\therefore E[(aX)^k]$ is exists.

Theorem 2.3.

Prove that $E[aX + b] = aE(X) + b$ where a, b are constant and $E(b) = b$

Proof.

Consider $E[aX + b] = E[aX] + E[b]$

$$= aE[X] + b$$

$\therefore E[aX + b] = aE[X] + b.$

Example 1.

The random variable X can take on two values 2 and 4 where $P(X = 2) = 0.2$ and $P(X = 4) = 0.8$. Find $E[X^2]$

Solution

$m_K = E[X^K] = \sum_l x_l^K P_l$

$E[X^2] = \sum_l x_l^2 P_l$

$$= (2)^2 P(X = 2) + 4^2 P(X = 2)$$

$$= 4(0.2) + 16(0.8)$$

$$= 0.8 + 12.8 = 13.6$$

$$m_2 = 13.6$$

Example 2.

The random variable X has the normal distribution with density $f(x) = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$. Find $E(X^2)$.

Solution.

$$m_K = E(X^K) = \int_{-\infty}^{\infty} x^K\, f(x)dx$$

$$m_2 = E(X^2) = \int_{-\infty}^{\infty} x^2 \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\, dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2\, e^{-\frac{x^2}{2}}\, dx$$

$$u = -x, dv = -xe^{-\frac{x^2}{2}} dx$$

$$du = -dx, v = e^{-\frac{x^2}{2}}$$

$$= \frac{1}{\sqrt{2\pi}} \left[\left(-e^{-\frac{x^2}{2}} x \right)_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx \right]$$

$$= \frac{1}{\sqrt{2\pi}} \left[0 + \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx \right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}\, dx$$

$$= 1$$

$m_2 = E(X^2) = 1$

Definition.

$E[(X-c)^k]$ where $c$is an arbitrary constant is called the ***moment of order k*** with respect to the point c.

Definition.

Moments with respect to the expected value, that is, with respect to the point $c = m_1 = E(X)$ are called ***central moments***.

We denoted it by $\mu_K = E\left[\left(X - E(X)\right)^K\right]$.

Moment with respect to the point $c = 0$ are called ***ordinary moments***.

Moment about mean

$\mu_K = E((X - E(X)^K)$

$\mu_1 = E\big(X - E(X)\big)$

$= E(X - m_1)$

$= E(X) - m_1 = m_1 - m_1$

$\mu_1 = 0$

$\mu_2 = E\left(\left(X - E(X)\right)^2\right)$

$= E((X - m_1)^2)$

$= E(X^2 - m_1^2 - 2m_1X)$

$= E(X^2) + m_1^2 - 2m_1E(X)$

$$= m_2 + m_1^2 - 2m_1^2$$

$$\mu_2 = m_2 - m_1^2$$

$$\mu_3 = E\left(\left(X - E(X)\right)^3\right)$$

$$= E((X - m_1)^3)$$

$$= E[X^3 - m_1^3 - 3X^2 m_1 + 3X m_1^2]$$

$$= E(X^3) - m_1^3 - 3E(X^2)m_1 + 3E(X)m_1^2$$

$$= m_3 - m_1^3 - 3m_2 m_1 + 3m_1^3 = m_3 + 2m_1^3 - 2m_1 m_2$$

Note

When $k = 2$,

The moment $E[(X - c)^2]$ is called the mean quadratic deviation of the random variable X from the point C.

Definition.

The central moment of the second order $\mu_2 = m_2 - m_1^2$ is called the ***variance***.

It is denoted by $D^2(X)\ or\ \sigma^2$.

Definition.

Square root of the variation is called the ***standard deviation***

$$i.e., S.D = \sqrt{\sigma^2} = \sigma$$

Example 5.

Compute mean and variation of Binomial distribution. The random variable X takes on the values $r = 0,1,2,\ldots\ldots,n$ with $P(X = r) = \binom{n}{r} P^r q^{n-r}$

Solution.

Given $P(X = r) = \binom{n}{r} P^r (1-p)^{n-r}$

$$E(X) = \sum_r x^r p^r$$

$$E(X) = \sum_{r=0}^{n} r \binom{n}{r} P^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} r \binom{n}{r} P^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} r \frac{n!}{r!(n-r)!} P^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} r \frac{n(n-1)!}{r(r-1)!(n-r)!} P\, P^r (1-p)^{n-r}$$

$$= np \sum_{r=0}^{n} \frac{(n-1)!}{(r-1)!\big(n-1-(r-1)\big)!} P^{r-1} (1-p)^{\left(n-1-(r-1)\right)}$$

$$= np \sum_{k=0}^{n-1} r \binom{n-1}{K} P^K (1-p)^{n-1-K}$$

$$= nP\big(P + (1P)\big)^{n-1}$$

$$= nP(P + 1 - P)^{n-1}$$

$$E(X) = nP$$

$$E(X^2) = \sum_{r=0}^{n} r^2\, P(X = r)$$

$$= \sum_{r=0}^{n} r^2 \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} \{r^2 - r + r\} \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} \{r\,(r-1) + r\} \frac{n!}{r!(n-r)!} p^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} (r(r-1) + r) \frac{n(n-1)}{r(r-1)} \frac{(n-2)!}{(r-2)!(n-2-(r-2)!} p^r (1-p)^{n-r}$$

$$= \sum_{r=0}^{n} r\,(r-1) \frac{n(n-1)}{r(r-1)} \frac{(n-2)!}{(r-2)!\big((n-2)-(r-2)\big)!} P^r q^{n-r} +$$

$$\sum_{r=0}^{n} r \frac{n(n-1)}{r(r-1)} \frac{(n-2)!}{(r-2)!\big((n-2)-(r-2)\big)!} P^r q^{n-r}$$

$$= n(n-1) \sum_{r=2}^{n} \binom{n-2}{r-2} p^{r-2+2} q^{(n-2)-(r-2)} +$$

$$\sum_{r=0}^{n} n \frac{(n-1)!}{(r-1)!(n-1)!-(r-1)!} p^{r-1+1} q^{(n-1)-(r-1)}$$

$$= n(n-1)p^2 \sum_{r=2}^{n} \binom{n-2}{r-2} p^{r-2} q^{(n-2)-(r-2)} +$$

$$np \sum_{r=1}^{n} \binom{n-1}{r-1} p^{r-1} q^{(n-1)-(r-1)}$$

$$= n(n-1)p^2 (p+q)^{n-2} + nP\,(p+q)^{n-1}$$

$$= n(n-1)p^2 (p+1-P)^{n-2} + np\,(P+1-P)^{n-1}$$

$E(X^2) = n(n-1)p^2 + np$

Variance $\sigma^2 = \mu_2 = m_2 - m_1^2$

$D^2(X) = n(n-1)p^2 + np - (np)^2$

$$= n^2p^2 - np^2 + np - n^2p^2$$

$$= np - np^2$$

$$= np(1-p) = npq$$

Example 6. Find μ_4 and m_4.

Suppose $n = 2, p = \frac{1}{2}$

Then $E(X) = np = 2 \times \frac{1}{2} = 1$

$\sigma^2 = npq = 2 \times \frac{1}{2} \times \left(1 - \frac{1}{2}\right)$

$= \frac{1}{2}$

$E(X) = 1, \sigma^2 = \frac{1}{2}$.

Example 7.

Let $n = 3, p = \frac{1}{3}$

$E(X) = np = 3 \times \frac{1}{3} = 1$

$\sigma^2 = npq = 3 \times \frac{1}{3} \times \left(1 - \frac{1}{3}\right)$

$= \frac{2}{3}$

Property of the variance

1. For every $c \neq m_1, D^2(X) < E[(X - X)^2]$.

   Proof.

   $E[(X - c)^2] = E[([X - m_1] + [m_1 - c])^2]$

   $= E[(X - m_1)^2 + (m_1 - c)^2 + 2(X - m_1)(m_1 - c)]$

   $= E[(X - m_1)^2] + (m_1 - c)^2 + 2(m_1 - c)E[(X - m_1)]$

   $= D^2(X) + (m_1 - c)^2 + 2(m_1 - c)(E(X) - m_1)$

   $= D^2(X) + (m_1 - c)^2$

   $\therefore E[(X - c)^2] = D^2(X) + (m_1 - c)^2$

   $D^2(X) < E[(X - c)^2]$

2. Find the variance of the liner function of the random variable $Y = X + C$ where c is constant.

Proof.

Let $Y = X + C$ where C is constant

$$D^2(Y) = E\left[\left(Y - E(Y)\right)^2\right]$$

$$= E\left[\left(X + C - E(X + X)\right)^2\right]$$

$$= E[(X + C - E(X) - C)^2]$$

$$= E\left[\left(X - E(X)\right)^2\right] = D^2(X)$$

$\therefore D^2(Y) = D^2(X)$

3. Find the variance of the random variable $Y = aX + b$ where a and b are constant.

Proof.

Let $Y = aX + b$

$$D^2(Y) = E(Y^2) - \left(E(Y)\right)^2$$

$$= E((aX + b)^2) - \left(E(aX + b)\right)^2$$

$$= E(a^2X^2 + b^2 + 2abX) - (aE(X) + b)^2$$

$$= a^2E(X^2) + b^2 + 2abE(X) - \left(a^2E(X)\right)^2 + b^2 + 2abE(X))$$

$$= a^2E(X^2) + b^2 + 2abE(X) - a^2\left(E(X)\right)^2 - b^2 - 2abE(X)$$

$$= a^2\left[E(X^2) - \left(E(X)\right)^2\right]$$

$$= a^2D^2(X)$$

$\therefore D^2(Y) = a^2D^2(X)$.

Definition

A random variable X for which $E(X) = 0, D^2(X) = 1$ is called the ***standardized random variable***.

Problem 2.1.

If X is the random variable with $E(X) = m_1$ & standard deviation σ, the random variable Y defined as $Y = \frac{X-m_1}{\sigma}$. Prove that Y is a standardized random variable.

Solution.

Let $Y = \frac{X-m_1}{\sigma}$

$$E(Y) = E\left(\frac{X-m_1}{\sigma}\right)$$

$$= \frac{1}{\sigma} E(X - m_1) = \frac{1}{\sigma}[E(X) - m_1] = 0$$

$$D^2(Y) = \sigma^2 = m_2 - m_1^2$$

$$= E[Y^2] - [E(Y)]^2 = E\left[\left(\frac{X-m_1}{\sigma}\right)^2\right] = 0$$

$$= \frac{E[X^2+m_1^2-2Xm_1]}{\sigma^2}$$

$$= \frac{1}{\sigma^2}[E(X^2) + m_1^2 - 2m_1E(X)]$$

$$= \frac{1}{\sigma^2}[m_2 + m_1^2 - 2m_1^2] = \frac{1}{\sigma^2}[m_2 - m_1^2]$$

$$= \frac{1}{\sigma^2} \times \sigma^2$$

$$D^2(Y) = 1$$

$\therefore E(X) = 0$ and $D^2(Y) = 1$

$\Rightarrow Y$ is the standardized random variable.

Remark.

The standard deviation, the mean deviation can serve as a measure of dispersssion.

i) If the variable X is of the discrete type with jumps points x_i

$$\sum_{-\infty<x_1<\infty} P_i|x_i - m_1|$$

ii) If X is the random variable of the continuous type $\int_{-\infty}^{\infty}|x - m_1|\, f(x)dx$.

Definition.

The ratio of the standard deviation to the expected value is called the ***coefficient of variation***. This ratio is denoted by $v = \frac{\sigma}{m_1}$

If $E(X) = m_1 = 1$ then $v = \sigma$

i.e., coefficient of variation = standard deviation

i.e., the coefficient of variation is a measure of dispersion if the expected value is the unit of measurement.

Example.

Find the co-efficient of variation in Binomial distribution.

Solution

$E(X) = np$ and $\sigma = \sqrt{npq}$

$$v = \frac{\sigma}{m_1} = \frac{\sqrt{npq}}{np} = \sqrt{\frac{q}{np}}$$

Definition.

The random variable X has a ***symmetric distribution*** if there exists a point a such that for every x the distribution $F(x)$ of X satisfy the equation.

$$F(a - x) = 1 - F(a + x) - P(X = a + x)$$

The point a is called the ***center of symmetry***.

In particular, If $a = 0 \; \forall x, F(x) = 1 - F(x) - P(X = x)$.

Definition.

If a random variable with symmetric distribution is of the continuous type, its density function $f(x)$ satisfies the equation (excluding the discontinuity points).

Definition.

If X is of discrete type, its jump points and their probabilities are placed symmetrically with respect to the center of symmetry.

Note

1. If the random variable has a symmetric distribution and its expected value exists, this expected value equals the center of symmetry.
2. For a symmetric distribution the central moments of r orders (if they exists) are equal to zero.
3. Since, in a symmetric distribution all the central limits of odd order equals zero, the value of every moments of odd order as a measure of asymmetric of the distribution

   For the measure of asymmetry we use the expression.

$$\gamma = \frac{\mu_3}{\sigma^3}$$

   This is called the co-efficient of skewness.

Theorem 2.3

Suppose that the moments $m_k (k = 1,2 \dots \dots)$ of a random variable X exists and the series $\sum_{k=1}^{\infty} \frac{m_k}{k!} r^k$ is absolutely convergent for some $r > 0$, then the set of moments $\{m_k\}$ uniquely determines the distribution function $F(x)$ of X.

Note

1. If for some of constant $\mu, |m_k| \leq \mu^k$ $(K = 1,2, \ldots\ldots)$ the distribution function $F(x)$ is uniquely convergent.

   If the set of all possible values of a random variable X is bounded from both sides, the set $\{m_k\}$ determines $F(x)$ uniquely.

Examples.

1. Let the random variable X takes on the values $x_k = \frac{2^k}{k^2}$ $(K = 1,2 \ldots\ldots)$ with probability is $p_k = \frac{1}{2^k}$.

   $E(X) = \sum_{k=1}^{\infty} x_k p_k = \sum_{k=1}^{\infty} \frac{2^k}{k^2} \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty$

   $E(X)$ exists.

   $E(X^2) = \sum_{k=1}^{\infty} x_k^2 \, P_k = \sum_{k=1}^{\infty} \frac{(2^k)^2}{k^4} \frac{1}{2^k} = \sum_{k=1}^{\infty} \frac{2^k}{k^4} = \infty$

2. Let a random variable Y take on zero with probability $\frac{1}{2}$ and the values $y_k = \frac{2^{k+1}}{k^2}$ with probabilities $P_k' = \frac{1}{2^{k+1}}$

   $E(Y) = \sum_{k=1}^{\infty} y_k \, P_k'$

   $= \sum_{k=1}^{\infty} \frac{2^{k+1}}{k^2} \frac{1}{2^{k+1}} = \sum_{k=1}^{\infty} \frac{1}{k^2} = E(X)$

   $E(Y^2) = \sum_{k=1}^{\infty} y_k^2 \; P_k' = \sum_{k=1}^{\infty} \frac{(2^{k+1})^2}{k^4} s \, . \frac{1}{2^{k+1}}$

   $= \sum_{k=1}^{\infty} \frac{2^{k+1}}{k^4} = \infty$

∴ The random variable X and Y have the same moments of the first order does not have moments of any order > 1.

2.3. The Chebyshev Inequality

Theorem 2.4.

If a random variable Y can take on only non negative values and has expected value $E(Y)$, then for an arbitrary positive number ks, $P(Y \geq k) \leq \frac{E(Y)}{k}$

Proof.

Suppose Y is of continuous type.

$$P(Y \geq K) = 1 - P(Y < K)$$

$$= \int_{-\infty}^{\infty} f(y)dx - \int_{-\infty}^{K} f(y)dy$$

$$= \int_{K}^{\infty} f(y)dy$$

Consider $E(Y) = \int_{-\infty}^{\infty} yf(y)dy$

$$= \int_{0}^{\infty} yf(y)dy$$

$$\geq \int_{K}^{\infty} y\, f(y)dy$$

$$\geq K \int_{K}^{\infty} f(y)dy$$

$$= KP(Y \geq K)$$

$$\therefore E(Y) \geq K\, P(Y \geq K)$$

$$\Rightarrow P(Y \geq K) \leq \frac{E(Y)}{K}$$

Corollary.

Let the random variable X have a distribution of probability about which we assume only that there is a finite variane σ^2 then for every $K > 0$, $P(|X - m_1| \geq K\sigma \leq \frac{1}{K^2}$

Proof.

Suppose the random variable X has the expected value $E(X) = m_1$ and a standard deviation σ

Consider the random variable $Y = (X - m_1)^2$

Clearly $Y \geq 0$

$$E(Y) = E((X - m_1)^2) = E(X^2 + m_1^2 - 2m_1X)$$

$$= E(X^2) + m_1^2 - 2m_1E(X)$$

$$= m_2 + m_1^2 - 2m_1^2 = m_2 - m_1^2$$

$$E(Y) = m_2 - m_1^2 = \sigma^2$$

$\therefore E(Y)$ is exists.

By the above theorem,

$$P(Y \geq K) \leq \frac{E(Y)}{K}$$

Put $E(Y) = \sigma^2$

And $K = K^2\sigma^2$ is a constant.

$$P((X - m_1)^2 \geq K^2\sigma^2) \leq \frac{\sigma^2}{K^2\sigma^2}$$

$$P((X - m_1)^2 \geq K^2\sigma^2) \leq \frac{1}{K^2}$$

$$P(|X - m_1| \geq K\sigma) \leq \frac{1}{K^2} \quad \dots\dots (1)$$

Note.

Equation (1) is valid for arbitrary random variables whose variance exists (second order moment exists)

Put $K = 3$ in equation (1),

$P(|X - m_1| \geq 3\sigma) \leq \frac{1}{9}$

If follows that in the class of random variables whose second order moment exists one can't obtain the better inequality than the Chebyshev's inequality.

Example 1.

The random variable X has the probability function

$$P(X = -K) = P(X = K) = \frac{1}{2K^2}, P(X = 0) = 1 - \frac{1}{K^2}$$

where K is some positive constants.

We have $E(X) = \sum_K x_K f(x)$

$= 0\, P(X = 0) + K\, P(X = K) + (-K)P(X = -K)$

$= K\left(\frac{1}{2k^2}\right) - K\left(\frac{1}{2K^2}\right) = 0$

$E(X^2) = \sum x_K^2\, P_K$

$= K^2\left(\frac{1}{2K^2}\right) + K^2\left(\frac{1}{2K^2}\right)$

$= 2K^2\left(\frac{1}{2K^2}\right) = 1$

$\sigma = m_2 - m_1^2 = 1 - 0 = 1$

We know that $= P(|X - m_1| \geq k\sigma) \leq \frac{1}{K^2}$

L.H.S $m_1 = 0, \sigma = 1$

$P(|X - m_1| \geq K\sigma) = P(|X| - K)$

$$= P(X = -K) + P(X = K) = \frac{1}{2K^2} + \frac{1}{2K^2} = \frac{2}{2K^2} = \frac{1}{K^2}$$

$P(X = -K) = P(X = K) = \frac{1}{2K^2}, P(X = 0) = 1 - \frac{1}{K^2}.$

2.4. Absolute Moment

Definition.

The expression $E(|X|^K)$ is called the **Absolute Moment of order k**. The absolute moment are denoted by β_k.

If X is a random variable of the discrete type with jump points x_i,

$$\beta_k = E(|X|^k = \sum_i |x|^k P_i$$

If X is a random variable of continuous type with density $f(x)$,

$$\beta_k = E(|X|^k) = \int_{-\infty}^{\infty} |x|^k f(x) dx$$

Remark.

1. The absolute moment of an even order equals the moment of the same order.

Theorem 2.5.(Lapunov inequality)

If for a random variable X the absolutely moment of order n exists, for arbitrary k $(k = 1,2, \dots \dots, n - 1)$ the following inequality is then true.

$$\beta_k^{\frac{1}{k}} \leq \beta_{k+1}^{\frac{1}{k+1}}$$

Proof.

Suppose that the random variable is of the continuous type is exists.

Let u and v be to arbitrary real numbers.

Consider the non-negative expression

$$\int_{-\infty}^{\infty}\left[u|x|^{\frac{k-1}{2}}+v|x|^{\frac{k+1}{2}}\right]^{2} f(x)dx = \int_{-\infty}^{\infty}[u^2|x|^{k-1}+v^2|x|^{k+1}+$$
$$2uv|x|^2 f(x)dx$$

$$=\int_{-\infty}^{\infty}u^2|x|^{k-1}f(x)dx+\int_{-\infty}^{\infty}u^2|x|^{k+1}f(x)dx+\int_{-\infty}^{\infty}2uv|x|^{k}f(x)dx$$

$$=u^2\beta_{k-1}+v^2\beta_{k+1}+2uv\beta_k$$

$$=u^2\beta_{k-1}+v^2\beta_{k+1}+2uv\beta_k\geq 0$$

i.e, $u^2\beta_{k-1}+2uv\beta_k+v^2\beta_{k+1}\geq 0$

we know that, the condition for the expression $ax^2+2hxy+by^2$ to be non-negative ∀ values of x and y is that $\begin{vmatrix} a & h \\ h & b \end{vmatrix}\geq 0$.

$\therefore \begin{vmatrix} \beta_{k-1} & \beta_k \\ \beta_k & \beta_{k+1} \end{vmatrix}\geq 0$

$\beta_{k-1}\beta_{k+1}-\beta_k^2\geq 0$

$\Rightarrow \beta_k^2\leq\beta_{k-1}\beta_{k+1} \dots\dots(1)$

Rising both the sides of equation (1) , to power k

$(\beta_K^2)^K\leq(\beta_{K-1}\beta_{K+1})^K$

$\beta_k^{2k}\leq\beta_{k-1}^k\beta_{k+1}^k \dots\dots(2)$

Put $k=1,2,\dots.n-1 \quad in\ (2)$

$\beta_1^2\leq\ \beta_0\beta_2$

$\beta_2^4\leq\beta_1^2\beta_3^2$

$\beta_3^6\leq\beta_2^3\beta_4^3$

$\beta_4^8\leq\beta_3^4\beta_5^4$

$\beta_5^{10} \leq \beta_4^5 \beta_6^5$

.

.

$\beta_{n-3}^{2(n-3)} \leq \beta_{n-4}^{n-3} \beta_{n-2}^{n-3}$

$\beta_{n-2}^{2(n-2)} \leq \beta_{n-3}^{n-2}\ \beta_{n-1}^{n-2}$

$\beta_{n-1}^{2(n-1)} \leq \beta_{n-2}^{n-1} \beta_n^{n-1}$

Where $\beta_0 = \int_{-\infty}^{\infty} |x|^0 f(x)dx = \int_{-\infty}^{\infty} f(x)dx = 1$

Multiplying the k successive inequalities and $\beta_0 = 1$, we get, $(\beta_{n-1} \leq \beta_n^{n-1})$

In general, $\beta_k \leq \beta_{k+1}\ \ \forall k = 1,2, \dots, -1 \ \ \rightarrow (3)$

Raising both side of (3) to power $\frac{1}{k(k+1)}$

$(\beta_k^{k+1})^{\frac{1}{k(k+1)}} \leq (\beta_{k+1}^{k})^{\frac{1}{k(k+1)}}$

$(\beta_k)^{\frac{1}{k}} \leq (\beta_k)^{\frac{1}{k+1}}$

$\beta_k^{\frac{1}{k}} \leq \beta_{k+1}^{\frac{1}{k+1}}$

2.5. Order Parameters

Definition.

The value x satisfying the inequalities

$P(X \leq x) \geq \frac{1}{2}, P(X \geq x) \geq \frac{1}{2}\ \dots \dots (1)$

is called the ***median*** and is denoted by $x_{\frac{1}{2}}$ one is equivalent to the double inequality.

$$\frac{1}{2} - P(X = x) \leq F(x) \leq \frac{1}{2} \quad \rightarrow (2)$$

If $P(X = x) = 0$

In particular,

If the random variable X is of the continuous type the median is the number x satisfying the equivality $F(x) = \frac{1}{2} \quad \rightarrow (3)$

If many value of x satisfying inequalities (1) or (2) then each of them is called the median.

Example.

1. Suppose that the random variable X can take on the values 0 and 1, $P(X = 0) = \frac{1}{5}, P(X = 1) = \frac{4}{5}$. Then find the median.

   Solution.

   Let $x = 0$

   $P(X \geq 0) = P(X = 0) + P(X = 1) = \frac{1}{5} + \frac{4}{5} = 1 > \frac{1}{2}$

   $P(X \leq 0) = P(X = 0) = \frac{1}{5} \ngtr \frac{1}{2}$

   ∴ 0 is not a median point.

   Let $x = 1$

   $P(X \geq 1) = P(X = 1) = \frac{4}{5} > \frac{1}{2}$

   $P(X \leq 1) = P(X = 0) + P(X = 1) = \frac{1}{5} + \frac{4}{5} = 1 > \frac{1}{2}$

   ∴ 1 is a median point.

2. The random variable X is of the continuous type with density defined as

$$f(x) = \begin{cases} 0 & for\ x < 0 \\ \cos x & for\ 0 \le x \le \frac{\pi}{2} \\ 0 & for\ x > \frac{\pi}{2} \end{cases}$$

Solution.

We know that, if the random variable X is of the continuous type, the median is the number $x_{\frac{1}{2}}$ satisfying the inequality $F\left(x_{\frac{1}{2}}\right) = \frac{1}{2}$

i.e., $\int_{-\infty}^{x_{\frac{1}{2}}} f(x)dx = \frac{1}{2}$

$$\int_{x_{\frac{1}{2}}}^{0} f(x)dx + \int_{0}^{x_{\frac{1}{2}}} f(x)dx + \int_{x_{\frac{1}{2}}}^{\infty} f(x)dx = \frac{1}{2}$$

$$\int_{0}^{x_{\frac{1}{2}}} \cos x\, dx = \frac{1}{2}$$

$$(\sin x)_{0}^{x_{\frac{1}{2}}} = \frac{1}{2}$$

$$\sin x_{\frac{1}{2}} = \frac{1}{2}$$

$$x_{\frac{1}{2}} = \frac{\pi}{6}$$

$\therefore \frac{\pi}{6}$ is a median point

3. Suppose that the random variable X can take on 3 values $x_1 = -1, x_2 = 0, x_3 = 1$ with $P(X = -1) = P(X = 0) = \frac{1}{4}$ $P(X = 1) = \frac{1}{2}$. Find the median.

Solution.

Let $x = -1$

$$P(X \ge -1) = P(X = -1) + P(X = 0) + P(X = 1)$$

$$= \frac{1}{4} + \frac{1}{4} + \frac{1}{2} = 1 > \frac{1}{2}$$

$$P(X \le -1) = P(X = -1) = \frac{1}{4} \ngtr \frac{1}{2}$$

$\therefore -1$ is not a median point.

$$P(X \ge 0) = P(X = 0) + P(X = 1)$$

$$= \frac{1}{4} + \frac{1}{2} = \frac{3}{4} > \frac{1}{2}$$

$$P(X \leq 0) = P(X = -1) + P(X = 0) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \geq \frac{1}{2}$$

$\therefore$ 0 is a median point.

$P(X \geq 1) = P(X = 1) = \frac{1}{2} \geq \frac{1}{2}$

$P(X \leq 1) = P(X = -1) + P(X = 0) + P(X = 1)$

$$= 1 > \frac{1}{2}$$

$\therefore$ 1 is a median point.

Here each value x from the interval $(0,1)$ is the median.

Definition.

The median is a special case of the class of parameters called quantiles.

Definition.

The value x satisfy the inequalities

$$P(X \leq x) \geq p, P(X \geq x) \geq 1 - p \; (0 < p < 1) \quad \ldots\ldots (1)$$

is called the ***quantile of order p*** and is denoted by x_p.

(1)is equivalent to the double inequality

$$p - P(X = x) \leq F(x) \leq p \quad \ldots\ldots (2)$$

If $P\left(X = x_p\right) = 0$. In particular, if the random variable X is of the continuous type.

$(2) \Rightarrow p \leq F(x) \leq p$

$\therefore F(x) = p$

The quantile of order p is the number satisfy the equation $F(x) = p \rightarrow \; (3)$

If many numbers x satisfy one or two each of them is then called the quantile of order p.

Definition

Quantiles and functions of them are called ***order parameters***.

Example.

1. Suppose that the random variable X has the normal distribution with density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Find the point x for which $F(x) = 0.1$.

   Solution.

   $x_{\frac{1}{10}} \cong 1.28$

Remark

1. Some simple functions of the quantiles may also serve as measure of dispersion.
2. The semi-inter quartile range defined as $\frac{1}{2}\left(x_{\frac{3}{2}} - x_{\frac{1}{4}}\right)$
3. If the set of all possible values of a random variable is bounded from both sides, there exists finite upper and lower bounds of the values taken by this random variables.
4. If a and b are the lower and upper bounds of the values taken on by the random variable the range is defined by $d = b - a$

   In example -2, the range equals $\frac{\pi}{2}$, in example – 3 it equals 2.

2.6. Moments of Random Vector

Let ordered pair of (X, Y) be a two-dimensional random variable. Consider the single valued functions $g(X, Y)$ of (X, Y).

Definition.

Let order pair of (X, Y) be a 2-dimensional random variable of the discrete type with jump points (x_i, y_K) and jumps P_{iK}.

The series

$$E[g(X,Y)] = \sum_{i,K} P_{iK}\, g(x_i\,, y_K)$$

is called the ***expected value*** of if the following inequality is satisfied:

$\sum_{i,k} P_{ik} |g(x_i\,, y_k)| < \infty$.

Definition.

Let (X, Y) be a random variable of the continuous type with density $f(x, y)$. Let $g(x, y)$ be Riemann Integrable.

$$E[g(X,Y)] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f(x,y) dx\, dy < \infty$$

Exercise

1. Express the central moment μ_4 as the function of the ordinary moments m_1, m_2, m_3, m_4.
2. Express the ordinary moment m_4 as the function of the central moments $\mu_1, \mu_2, \mu_3, \mu_4$
3. Show that if X_1 and X_2 are independent and have same distribution, $Y = X_1 - X_2$ has a symmetric distribution.

Definition.

The expected value of the function $g(X, Y) = X^l Y^n$

i.e., $m_{ln} = E(X^l . Y^n)$

where l and n are non-negative integer, is called the ***moment of order*** $\boldsymbol{l+n}$ of the random variable (X, Y).

Thus, if (X, Y) is random variable of the discrete type with jump points (x_i, y_k) and jumps P_{ik},

$$m_{ln} = \sum_{i,k} P_{ik}\, x_i^l y_k^n$$

If (X, Y) is a random variable of the continuous type with density function $f(x, y)$,

$$m_{ln} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^l y^n f(x, y) dx\, dy$$

Remark

1. Two moments of the first order exist

   $m_{10} = E(X^1Y^0)$

   $\quad = E(X)$

   $m_{01} = E(X^0Y^1)$

   $\quad = E(Y)$

   $\therefore m_{10} = E(X)$ and $m_{01} = E(Y)$
2. Three moments of the second order exists.

   $m_{20} = E(X^2Y^0) = E(X^2)$

   $m_{02} = E(X^0Y^2) = E(Y^2)$

   $m_{11} = E(XY)$

   $\therefore m_{20} = E(X^2), m_{02} = E(Y^2), m_{11} = E(XY)$.

Definition.

The central moment is denoted by μ_{ln}

$$\mu_{ln} = E[(X - m_{10})^l (Y - m_{01})^n]$$

Central moment of order 1:

$$\mu_{10} = E[(X - m_{10})^1 (Y - m_{01})^0]$$

$$= E[(X - m_{10})]$$

$$= E[X - E(X)]$$

$$= E(X) - E(X) = 0$$

$$\mu_{10} = 0$$

Similarly,

$$\mu_{01} = E[(Y - m_{01})] = 0$$

Central moment of order 2:

$$\mu_{20} = E[(X - m_{10})^2]$$

$$= E[X^2 + m_{10}^2 - 2m_{10}X]$$

$$= E(X^2) + m_{10}^2 - 2m_{10}E(X)$$

$$= E(X^2) + m_{10}^2 - 2m_{10}^2 = E(X^2) - m_{10}^2$$

$$\mu_{20} = m_{20} - m_{10}^2 = \sigma_1^2$$

Similarly, $\mu_{02} = m_{02} - m_{01}^2 = \sigma_2^2$

Where σ_1 and σ_2 are the standard deviation of the random variables X and Y respectively.

$\mu_{11} = E[(X - m_{10})(Y - m_{01})]$ is called co-variance and is denoted by $cov(X, Y) = E[(X - m_{10})(Y - m_{01})]$.

Relation between the ordinary and centre moments

Let (X, Y) be the random variables and the expected values of X and Y exists. Then,

$\mu_{20} = m_{20} - m_{10}^2$

$\mu_{02} = m_{02} - m_{01}^2$

$\mu_{11} = m_{11} - m_{10}m_{01}$

Proof.

$\mu_{20} = E[(X - m_{10})^2]$

$= E[X^2 + m_{10}^2 - 2m_{10}X\ ]$

$= E(X^2) + m_{10}^2 - 2m_{10}E(X)$

$= E(X^2) + m_{10}^2 - 2m_{10}^2 = E(X^2) - m_{10}^2$

$\therefore \mu_{20} = m_{20} - m_{10}^2 = \sigma_1^2$

$\mu_{02} = E[(Y - m_{01})^2]$

$= E[Y^2 + m_{01}^2 - 2Ym_{01}]$

$= E(Y^2) + m_{01}^2 - 2m_{01}^2 = E(Y^2) - m_{01}^2$

$\therefore \mu_{02} = m_{02} - m_{01}^2 = \sigma_2^2$

$\mu_{11} = E[(X - m_{10})(Y - m_{01})]$

$= E[XY - m_{10}Y - m_{01}X + m_{10}m_{01}]$

$= E(XY) - m_{10}E(Y) - m_{01}E(X) + m_{10}m_{01}$

$= E(XY) - m_{10}m_{01} - m_{01}m_{10} + m_{10}m_{01}$

$\mu_{11} = m_{11} - m_{01}m_{10}$

$\therefore cov(X,Y) = m_{11} - m_{01}m_{10}$

$cov(X,Y) = E(XY) - E(X)E(Y).$

Theorem 2.6.

The expected value of the sum of an arbitrary finite number of random variable, whose expected value exists, equals the sum of the expected values.

Proof.

We prove this theorem by induction to an arbitrary finite number of random variables.

Let the expected value $E(X)$ and $E(Y)$ are exists.

Let (X, Y) be a random variable of the discrete type with jump points (x_i, y_k) and jumps P_{iK}.

Let $Z = X + Y$

$$E(Z) = E(X + Y)$$

$$= \sum_{i,K} P_{iK}(x_i \, , y_K)$$

$$= \sum_{i,K}(P_{iK}x_i \, , P_{iK}y_K)$$

$$= \sum_{i,K} P_{iK}x_i + \sum_{i,K} P_{iK}\, y_K$$

$$= \sum_i P_{i.}x_i + \sum_{i,K} P_{.K}\, x_K$$

$$= E(X) + E(Y)$$

$$\therefore E(X + Y) = E(X) + E(Y)$$

$$\sum_{i,K} P_{iK}|x_i \, , y_K| \leq \sum_{i,K} P_{iK}(|x_i| \, , |y_K|)$$

$$= \sum_{i,K} P_{iK}|x_i| + \sum_{i,K} P_{iK}\, |y_K|$$

$$= \sum_i P_{i.}|x_i\,| + \sum_K P_{.K}|y_K|$$

$$< \infty \quad [\because E(X) \; and \; E(Y) < \infty$$

$\therefore E(X + Y)$ is exists.

Suppose the random variable (X,Y) is of continuous type with density $f(x,y)$

Since $E(X)$ and $E(Y)$ exists.

$\therefore E(X+Y)$ is exists.

$E(X+V)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)dx\,dy$

$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\big(x\,f(x,y)+yf(x,y)\big)dx\,dy$

$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xf(x,y)dx\,dy+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}yf(x,y)dx\,dy$

$=\int_{-\infty}^{\infty}x[\int_{-\infty}^{\infty}f(x,y)dy]dx+\int_{-\infty}^{\infty}y[\int_{-\infty}^{\infty}f(x,y)dx\,]dy$

$=\int_{-\infty}^{\infty}xf_1(x)dx+\int_{-\infty}^{\infty}y\,f_2(y)dy$

$E(X+Y)=E(X)+E(Y)$

$\therefore$ The result is true for two random variable.

Assume that $X_1,X_2,\ldots\ldots X_{n-1}$ be a random variable and the expected values $E(X_1),E(X_2),\ldots..E(X_{n-1})$ exist. Such that $E(X_1,X_2,\ldots\ldots X_{n-1})=E(X_1),E(X_2),\ldots..E(X_{n-1})$

Let $X_1,X_2,\ldots\ldots X_n$ be random variable such that $E(X_1),E(X_2),\ldots..E(X_n)$ exists

To prove:

$E(X_1,X_2,\ldots\ldots X_n)=E(X_1),E(X_2),\ldots..E(X_n)$

We know that $E(X_1,X_2,\ldots\ldots X_{n-1})$ and $E(X_n)$ exist

$\Rightarrow E(X_1,X_2,\ldots\ldots X_{n-1}+X_n)$ is exists.

$$E(X_1,X_2,\ldots\ldots X_{n-1}+X_n)=E\big((X_1,X_2,\ldots\ldots X_{n-1})+X_n\big)$$

$$=E(X_1,X_2,\ldots\ldots X_{n-1})+E(X_n)$$

$$=E(X_1)+E(X_n)+\cdots\ldots.E(X_n)$$

∴ The result is true for n number of random variable.

∴ By induction hypothesis, the result is true for arbitrary finite number of random variables.

Theorem 2.7.

The expected value of the product of the arbitrary finite number of independent random variable, whose expected values exists equals the product of the expected values of these variables.

Proof.

We prove this theorem by induction to an arbitrary finite number of independent random variables.

Let (X, Y) be a random variable of the discrete type such that X and Y are independent variables.

And $E(X)$ and $E(Y)$ exist.

$E(X, Y) = \sum_{i,k} P_{ik} x_i \, , y_k$

$= \sum_{i,k} P_{i.} \, P_{.k} x_i \, y_k$

$= \sum_i P_{i.} x_i \; \sum_k P_{.k} \, y_k$

$= E(X)E(Y)$

∴ $E(XY) = E(X)E(Y)$

$\sum_{i,K} P_{iK} |x_i \, , y_K| = \sum_{i,K} P_{i.} \, P_{.K} |x_i| |\, y_K| \; = \; \sum_i P_{i.} \, |x_i \,| \sum_K \; P_{.K} \, |y_K| < \infty$

∴ $E(XY)$ is exists.

If (X, Y) is of continuous type

$E(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xyf(x, y) dx \, dy$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy\, f_1(x)\, f_2(y) dx\, dy \quad [\because x \text{ and } y \text{ are independent}]$$

$$= \int_{-\infty}^{\infty} x f_1(x) dx \int_{-\infty}^{\infty} y\, f_2(y)\, dy$$

$$= E(X)E(Y)$$

$\therefore E(XY) = E(X)E(Y)$

$\therefore$ The result is true for two independent random variable.

Suppose $X_1, X_2, \dots \dots X_{n-1}$ are independent random variables such that $E(X_1), E(X_2), \dots \dots E(X_{n-1})$

Let $X_1, X_2, \dots \dots X_n$ are independent random variable.

$$E(X_1, X_2, \dots \dots X_n) = E\big((X_1, X_2, \dots \dots X_{n-1})X_n\big)$$

$$= E(X_1, X_2, \dots \dots X_{n-1})E(X_n)$$

$$= E(X_1), E(X_2), \dots \dots E(X_{n-1})E(X_n)$$

$\therefore$ The result is true for n-independent random variable.

Hence result is true for an arbitrary finite number of independent random variable.

Corollary.

The covariance of two independent random variable equals to 0.

Proof.

Since X and Y are independent.

By above theorem, $E(XY) = E(X)E(Y)$

$$\therefore cov\ (X_i Y) = E(XY) - E(X)E(Y) = 0.$$

Problem 1.

Let X and Y are two random variable with variance $D^2(X)$ and $D^2(Y)$. Let $Z = X + Y$. Find the variance of Z.

Solution

Let $Z = X + Y$

$$D^2(Z) = D^2(X+Y)$$
$$= E[(X+Y)^2] - [E(X+Y)]^2$$
$$= E[X^2 + Y^2 + 2XY] - [E(X) + E(Y)]^2$$

$$= E(X^2) + E(Y^2) + 2E(XY) - \left[\left(E(X)\right)^2 + \left(E(Y)\right)^2 + 2E(X)E(Y)\right]$$

$$= E(X^2) - [E(X)]^2 + E(Y^2) - [E(Y)]^2 + 2E(XY) - 2E(X)E(Y)$$

$$D^2(X+Y) = D^2(X) + D^2(Y) + 2E(XY) - 2E(X)E(Y)$$

If X and Y are independent random variable.

$$E(X,Y) = E(X)E(Y)$$

$$\therefore D^2(X+Y) = D^2(X) + D^2(Y)$$

Theorem 2.8.

The variance of the sum of an arbitrary finite number of independent random variable choose variance exist, equals the sum of the variance.

Proof. as in theorem 2.6.2

Conditional Expected value

Definition.

Let (X, Y) be a random variable of the discrete type with jump points (x_i, y_k) and jumps P_{ik}. Then the conditional expected value of the random variable Y^l under the condition $X = x_i$ is

$$E[Y^l|X = x_i] = \sum_k y_k^l \frac{p_{iK}}{p_{i\cdot}}$$

Similarly, the conditional expected value of X^l given $Y = y_k$ is

$$E[X^l|Y = y_k] = \sum_i x_i^l \frac{p_{ik}}{p_{.k}}$$

Definition.

Let (X, Y) be two dimensional random variable of the continuous type with density function $f(x, y)$ and conditional density $f_1(x)$ and $f_2(y)$ exist, we obtain

$$E(Y^l|X = x) = \int_{-\infty}^{\infty} y^l \frac{f(x, y)}{f_1(x)} dy$$

$$E(X^l|Y = y) = \int_{-\infty}^{\infty} x^l \frac{f(x, y)}{f_2(x)} dx$$

Remark.

1. For every subset S of the set of jump points x_i of X,

$$E(Y^l + X \in S) = \sum_k y_k^l P(Y = y_k|X \in S)$$

$$= \sum_k y_k^l \frac{\sum_{x\in S} p_{ik}}{\sum_{x\in S} p_{i\cdot}}$$

$$\therefore \sum_{x\in S} p_{i\cdot} \frac{\sum_k \frac{y_k^l P_{ik}}{p_{i\cdot}}}{\sum_{x_i\in S} p_{i\cdot}} = \sum_{x_i\in S} \left(\frac{p_{i.}}{\sum_{x_i\in S} p_{i\cdot}}\right) E(y_k^l = x_i)$$

$$E(Y^l + X \in S) = E\left[E(Y^l|X = x_i)\middle|X \in S\right] \ldots\ldots (1)$$

$$E(Y^l) = \sum_i \sum_k y_k^l P_{ik}$$

$$= \sum_i p_i . E(Y^l|X = x_i)$$

$$E(Y^l) = E\left[E(Y^l|X)\right] \quad \ldots\ldots (2)$$

Similarly,

For a random variable (X, Y) of the continuous type for ever Borel set S on the real axis for which $P(X \in S) > 0$,

$$E(Y^l|X \in S) = \int_{-\infty}^{\infty} y^l \int_S \frac{f(x,y)}{P(X \in S)} dxdy$$

$$= \int_S \frac{f_1(x)}{P(X \in S)} \int_{-\infty}^{\infty} y^l \frac{f(x,y)}{f_1(x)} dy\, dx$$

$$E(Y^l|X \in S) = \int_S \frac{f_1(x)}{P(X \in S)} E(Y^l|X = x)\, dx \quad \ldots\ldots (3)$$

Put $S = (-\infty, \infty)$ we obtain,

$$E(Y^l) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^l f(x,y) dx\, dy$$

$$= \int_{-\infty}^{\infty} f_1(x) E(Y^l|X = x) dx$$

$$= E\left[E\left(Y^l\middle|x\right)\right]$$

$$E(Y^l) = E\left[E\left(Y^l\middle|x\right)\right] \quad \ldots\ldots (4)$$

Formulas (1)And (3) may be written as,

$$E(Y^l|X \in S) = E\left[E\left(Y^l\middle|X = x\right)\middle|X \in S\right] \quad \ldots\ldots (5)$$

(2) and (4) may be written as,

$$E(Y^l) = E\left[E\left(Y^l\middle|x\right)\right] \quad \ldots\ldots (6)$$

2. Consider a random variable $(X_1, X_2, \ldots\ldots X_m)$ of the continuous type and suppose that the density functions $f(x_1, x_2, \ldots\ldots x_m)$ is everywhere continuous and the density of the marginal distribution

$$\int_{-\infty}^{\infty} f(x_1, x_2, \ldots\ldots x_m) dx_1 = g(x_2, \ldots\ldots x_m)$$

of the random variable $(X_2, \ldots\ldots X_m)$ is everywhere continuous and positive.

Conditional moments of order $l(l = 1,2, \dots\dots)$ is

$$E\left(X_1^l \middle| X_2 = x_2, \dots \dots x_m = x_m\right) = \int_{-\infty}^{\infty} x_1^l \frac{f(x_1,x_2,\dots\dots x_m)}{g(x_2,\dots\dots x_m)}\, dx_1.$$

Definition.

The formula for the ***coefficient of correlation*** is the following:

$$\rho = \frac{E[(X-m_{10})(Y-m_{01})}{\sqrt{E[(X-m_{10})^2}\sqrt{E[(Y-m_{01})^2}} = \frac{\mu_{11}}{\sigma_1\sigma_2}$$

Theorem 2.9.

The co-efficient of correlation satisfies the double inequality $-1 \leq \rho \leq 1$.

Proof.

For arbitrary real numbers t and u, consider the non-negative expression

$E\{[t(X - m_{10}) + u(Y - m_{01})]^2\}$

$= E\{t^2(X - m_{10})^2 + u^2(Y - m_{01})^2 + 2tu(X - m_{10})(Y - m_{01})\}$

$= E[t^2(X^2 + m_{10}^2 - 2Xm_{10})] + E[u^2(Y^2 + m_{01}^2 - 2Ym_{01})] + E[2tu(XY - m_{01}X - m_{10}Y + m_{10}m_{01})]$

$= t^2[E(X^2) + m_{10}^2 - 2m_{10}E(X)] + u^2[E(Y^2) + m_{01} - 2m_{01}E(Y)] + 2tuE(XY) - m_{01}E(X) - m_{10}E(Y) + m_{10}m_{01}$

$= t^2[m_{20} + m_{10}^2 - 2m_{10}m_{10}] + u^2[m_{02} + m_{01}^2 - 2m_{01}m_{01}] + 2tu[m_{11} - m_{01}m_{10} - m_{10}m_{01} + m_{10}m_{01}]$

$= t^2[m_{20} - m_{10}^2] + u^2[m_{02} - m_{01}^2] + 2tu[m_{11} - m_{10}^2 - m_{10}m_{01} + m_{10}^2]$

$= t^2\, \sigma_1^2 + u^2\sigma_2^2 + 2tu\, \mu_{11}$

$\therefore E\{[t(X - m_{10}) + u(Y - m_{01})]^2\}$

$$= t^2\sigma_1^2 + u^2\sigma_2^2 + 2tu\,\mu_{11} \dots \dots (1)$$

$L.H.S$ of (1), is always non-negative, we must have

$\therefore \mu_{11}^2 - \sigma_1^2\sigma_2^2 \le 0$ $\quad$ $[\because \begin{vmatrix} a & h \\ h & b \end{vmatrix} \ge 0 \quad \begin{vmatrix} \sigma_1^2 & \mu_{11} \\ \mu_{11} & \sigma_2^2 \end{vmatrix} \ge 0$

$\Rightarrow \sigma_1^2\sigma_2^2 - \mu_{11}^2 \ge 0$

$\Rightarrow \mu_{11}^2 \le \sigma_1^2\sigma_2^2$

$\Rightarrow \mu_{11} \le \pm\sigma_1\sigma_2$

$\Rightarrow -\sigma_1\sigma_2 \le \mu_{11} \le \sigma_1\sigma_2$

$\Rightarrow -1 \le \frac{\mu_{11}}{\sigma_1\sigma_2} \le 1$

$\Rightarrow -1 \le \rho \le 1$

Note.

1. If the random variable X and Y are independent $cov(X, Y) = \mu_{11} = 0$

   $\therefore \rho = 0$

   If the random variable X and Y are independent, then $\rho = 0$
2. But the converse is not true

   If $\rho = 0$ we say that X and Y are uncorrelation.

Theorem 2.10

The equality $\rho^2 = 1$ is the necessary and sufficient condition for the relation $P(Y = aX + b) = 1$ to hold.

Proof.

Suppose that $P(Y = aX + b) = 1$

i.e., $P(Y \ne aX + b) = 0$

We know that,

$m_{01} = E(Y)$

$= P(Y = aX + b)E(Y|Y = aX + b) + P(Y \neq aX + b)E(Y|Y \neq aX + b)$

$= E(Y|Y = aX + b) + 0$

$= E(aX + b)$

$= aE(X) + b$

$i.e., m_{01} = am_{10} + b \quad \dots\dots (1)$

$\sigma_2^2 = E[(Y - m_{01})^2]$

$= P[Y = aX + b]E[(Y - m_{01})^2|Y = aX + b]$
$\qquad + P(Y \neq aX + b)E[(Y - m_{01})^2|Y \neq aX + b]$

$= E[(Y - m_{01})^2|Y = aX + b]$

$= E[(aX + b - m_{01})^2]$

$= E[(aX + b - am_{10} - b)^2]$

$= E[a^2(X - m_{10})^2] = a^2E[(X - m_{10})^2]$

$i.e., \sigma_2^2 = a^2\sigma_1^2 \qquad \rightarrow (2)$

$\mu_{11} = E[(X - m_{10})(Y - m_{01})] = E[(X - m_{10})(aX + b - m_{01})]$

$= E[(X - m_{10})(aX + b - am_{10} - b)] \quad using\ (1)$

$= E[a(X - m_{10})^2]$

$\mu_{11} = aE[(X - m_{10})^2]$

$\mu_{11} = a\sigma_1^2$

$\rho = \frac{\mu_{11}}{\sigma_1\sigma_2} = \frac{a\sigma_1^2}{\sigma_1 a\sigma_1} = 1$

$\therefore \rho^2 = 1$

Conversely, Suppose that $\rho^2 = 1$

$\frac{\mu_{11}^2}{\sigma_1^2\sigma_2^2} = 1$

$\mu_{11}^2 = \sigma_1^2\sigma_2^2$

$\sigma_1^2\sigma_2^2 - \mu_{11}^2 = 0 \quad \rightarrow (2)$

$\mu_{20}\mu_{02} - \mu_{11}^2 = 0$

For arbitrary real number t and u consider the non-negative expression

$E\{[t(X - m_{10}) + u(Y - m_{01})]^2\}$

$= E\{t^2(X - m_{10})^2 + u(Y - m_{01})^2 + 2tu(X - m_{10})(Y - m_{01})\}$

$= E\{t^2[X^2 + m_{10}^2 - 2Xm_{10}]\} + E\{u^2[Y^2 + m_{01}^2 - 2Ym_{01}]\} + E\{2tu[XY - Xm_{01} - Ym_{10} + m_{10}m_{01}]\}$

$= \{t^2[E(X^2) + m_{10}^2 - 2m_{10}E(X)] + u^2[E(Y^2) + m_{01}^2 - 2m_{01}E(Y)] + 2tu[E(XY) - m_{01}E(X) - m_{10}E(Y) + m_{10}m_{01}]\}$

$= t^2[m_{20} + m_{10}^2 - 2m_{10}m_{10}] + u^2[m_{02} + m_{01}^2 - 2m_{01}m_{01}] + 2tu[m_{11} - m_{01}m_{10} - m_{10}m_{01} + m_{10}m_{01}]$

$= t^2[m_{20} - m_{10}]^2 + u^2[m_{02} - m_{01}^2] + 2tu[m_{11} - m_{01}m_{10}]$

$= t^2\sigma_1^2 + u^2\sigma_2^2 + 2tu\mu_{11} \quad \rightarrow (4)$

Since LHS of (4) is always non-negative.

$\sigma_1^2\sigma_2^2 - \mu_{11}^2 \geq 0$

i.e, $\mu_{20}\mu_{02} - \mu_{11}^2 \geq 0$

but (3) $\Rightarrow \mu_{20}\mu_{02} - \mu_{11}^2 = 0$

Then the quadratic form equation (4) takes on the value zero for some pair of values $t = t_0$ and $u = u_0$ where atleast one of the value t_0 and u_0 is not zero.

For these values t_0 and u_0 we have

$$E\{[t(X - m_{10}) + u(Y - m_{01})]^2\} = 0$$

This equation is satisfy only when we have the equation

$$P[t_0(X - m_{10}) + u_0(Y - m_{01}) = 0] = 1$$

Suppose that $u_0 \neq 0$

$$P\left[\frac{t_0X}{u_0} - \frac{t_0m_{10}}{u_0} + Y - \frac{u_0m_{01}}{u_0} = 0\right] = 1$$

$$P\left[Y = \frac{-t_0}{u_0}X + \frac{m_{01}u_0 + m_{10}t_o}{u_0}\right] = 1$$

i.e., $P[Y = aX + b] = 1$, where $a = \frac{-t_0}{u_0}$ and $b = \frac{m_{01}u_0 + m_{10}t_0}{u_0}$.

Definition.

Consider the n-dimensional random variables $(X_1, X_2, \dots \dots, X_n)$. Suppose that the variance $\sigma_i^2, (i = 1,2, \dots \dots, n)$ of the random variable X_i exists and are positive. Then the covariance of all pairs of these random variables are also exist.

Let λ_{ik} and P_{ik} be the co-variance and the co-efficient of correaltion of X_i and X_k respectively.

The symmetric matrix

$$M = \begin{bmatrix} \lambda_{11} & \lambda_{12} \dots \dots & \lambda_{1n} \\ \lambda_{21} & \lambda_{22} \dots \dots & \lambda_{2n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot \quad \dots \dots & \cdot \\ \lambda_{n1} & \lambda_{n2} & \lambda_{nn} \end{bmatrix}$$

is called the ***matrix of second order moments***. The determinant of the matrix M is denoted by $|M|$.

Theorem 2.11.

The probability that the random variable $X_1, X_2, \dots.. X_n$ whose variation exists, satisfy atleast one linear relation equals 1 iff $|M| = 0$

Proof.

By the previous theorem, If $\rho^2 = 1$ is a necessary and sufficient condition for the relation $P(Y = aX + b) = 1$ to hold.

$$\text{If } E\left\{\left[\sum_{i=1}^{n} t_i \left(X_i - E(X_i)\right)\right]^2\right\} = \sum_{i,k=1}^{n} \lambda_{ik}\, t_i t_k \geq 0$$

There are linear relation among $X_1, X_2, \dots.., X_n$ and by the definition we get $|M| = 0$

Conversely, if $|M| = 0$ then the whose mass of probability is concentrate on a hyper plane of dimension less than n.

i.e, $P(Y = aX + b) = 1$

∴ There exist a linear relation $X_1, X_2, \dots.., X_n$ among themselves.

Definition

If the components $X_1, X_2, \dots \dots, X_n$ of the random vector $(X_1, X_2, \dots.. X_n)$ satisfy atleast one linear relation with probability 1, then the distribution $(X_1, X_2, \dots \dots X_n)$ is ***degenerate***.

If the determinant $|M| \neq 0$, the distribution of $(X_1, X_2, \dots, X_n)$ is ***nondegenerate.***

Definition.

The determinant $|M|$ is called ***the generalized variance***.

Definition

The expression $\sqrt{|R|}$ is called the ***scattered coefficient,*** *the* R is the matrix of the correlation coefficient P_{ik}, taking $\rho_{ii} = 1$;

$$\therefore R = \begin{bmatrix} 1 & \rho_{12} \dots \dots & \rho_{1n} \\ \rho_{21} & 1 \dots \dots & \rho_{2n} \\ \rho_{n1} & \rho_{n2} \dots \dots & \rho_{nn} \end{bmatrix}$$

Remark.

We know that, $\rho_{ik} = \frac{\lambda_{ik}}{\sigma_i \sigma_k}$

Clearly $|M| = \sigma_1^2 \sigma_2^2 \dots \dots \sigma_n^2 |R|$

The matrix R is also symmetric and its determinant satisfies the relation $|R| \leq 1$.

2.7. Regression of First type

Let (X, Y) be a two-dimensional random variable of the discrete type with jump point (x_i, y_k), and jumps p_{ik} and let $P_i.$ and $P_{.k}$ denote the probabilities in the marginal distribution of X and Y respectively.

Consider the conditional expectations of X and Y denoted by $m_1(y_k)$ and $m_2(x_i)$ respectively. Thus

$$m_1(y_k) = E(X|Y = y_k) = \sum_i x_i \frac{P_{ik}}{P_{.k}} \rightarrow (1)$$

$$m_2(x_i) = E(Y|X = x_i) = \sum_k y_k \frac{P_{ik}}{P_{i.}} \rightarrow (2)$$

We obtain two collection of points in the plane (x, y) from (1) consist of points with co-ordinate $x = m_1(y_k), y = y_k$

From (2), consist of points with co-ordinates $x = x_i, y = m_2(x_i)$.

Definition.

Let (X, Y) be a two-dimensional random variable of the continuous type with density $f(x, y)$ and marginal densities $f_1(x)$ and $f_2(y)$. The conditional expectations $m_1(y)$ and $m_2(x)$ are

$$m_1(y) = E(X|Y = y) = \int_{-\infty}^{\infty} x \frac{f(x, y)}{f_2(y)} dx \quad \rightarrow (3)$$

$$m_2(y) = E(Y|X = x) = \int_{-\infty}^{\infty} y \frac{f(x, y)}{f_1(x)} dy \quad \rightarrow (4)$$

Again we obtain two collection of points in the plane (x, y) with the respective coordinates

From (3), $l = m_1(y), y$

From (4), $x, y = m_2(x)$

$x = x_i, y = m_2(x_i)$.

Example 1.

The random variable X and Y have the joint density given by the formula $f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 - 2xy + 2y^2}{2}\right)$ Find the correlation coefficient (or) Find the correlation coefficient of two-dimensional normal distribution.

Solution.

$f(x, y) = \frac{1}{2\pi} \exp\left(-\frac{x^2 - 2xy + 2y^2}{2}\right)$

This is a density function, since it is non-negative

And $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dx\,dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}e^{-\left(\frac{x^2-2xy+2y^2}{2}\right)}dx\,dy$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2}}dx\right)dy$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}(1)dy \qquad \left[\because t = x - y\; dt = dx\; \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2}}\boldsymbol{dx} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}dt\right]$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}dy$$

$$= 1$$

Clearly, $X \sim n(0,1)$ & $Y \sim \cap(0,1)$

$m_{10} = E(X^1Y^0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xf(x,y)dx\,dy$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x\,e^{-\left(\frac{(x-y)^2}{2}\right)}\boldsymbol{dx}\right)\boldsymbol{dy}$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} ye^{-\frac{y^2}{2}}dy$$

$m_{10} = 0$

$E(X^0Y^1) = m_{01} = \frac{1}{2\pi}\int_{-\infty}^{\infty} ye^{-x^2}$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} ye^{-\left(\frac{x^2-2xy+2y^2}{2}\right)}dy\,dx$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\,e^{-\left(\frac{x^2}{2}+\left(y-\frac{x}{2}\right)^2\right)}dy\,dx$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{x^2}{4}}\left(\int_{-\infty}^{\infty} ye^{-\left(y-\frac{x}{2}\right)^2}dy\right)dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4}} \left(\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} y e^{-\left(y-\frac{x}{2}\right)^2} \, dy \right) dx$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{x^2}{4}} \left(\frac{x}{2} \right) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{x}{2} e^{-\frac{x^2}{4}} dx = 0$$

$$\mu_{11} = m_{11} - m_{01} m_{10}$$

$$= m_{11}$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \, e^{-\left(\frac{y^2}{2}\right)} e^{-\frac{(x-y)^2}{2}} dx \, dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y \, e^{-\left(\frac{y^2}{2}\right)} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{-\frac{(x-y)^2}{2}} dx \right) dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y e^{-\left(\frac{y^2}{2}\right)} (y) dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy$$

$$= E(Y^2)$$

$$\mu_{11} = 1$$

$$\mu_{20} = m_{20} m_{10}^2$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^2 e^{-\frac{y^2}{2}} e^{-\frac{(x-y)^2}{2}} dx \, dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{(x-y)^2}{2}} dx \right] dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \left[\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (t^2 + y^2 + 2ty) e^{-\frac{t^2}{2}} \, dt \right]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \left[-\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} t^2 e^{-\frac{y^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{t^2}{2}} dt + \right.$$

$$\left. \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} 2ty e^{-\frac{t^2}{2}} dt \right]$$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} \left[\frac{2}{\sqrt{2\pi}} \times \frac{\sqrt{2\pi}}{2} + y^2 + 0\right] dy$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} [1 + y^2] dy$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy$

$= 1 + \frac{2}{\sqrt{2\pi}} \frac{\sqrt{2\pi}}{2} = 1 + 1 = 2$

$\mu_{02} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} e^{-\frac{(x-y)^2}{2}} \, dx \, dy$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-y)^2}{2}} dx\right) dy$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy$

$= \frac{1}{\sqrt{2\pi}} 2 \int_{0}^{\infty} y^2 e^{-\frac{y^2}{2}} dy$

$\mu_{02} = \frac{2}{\sqrt{2\pi}} \frac{\sqrt{2\pi}}{2} = 1$

$\rho = \frac{\mu_{11}}{\sigma_1 \sigma_2}$

$\mu_{20} = \sigma_1^2 = 2, \mu_{02} = \sigma_1^2 = 1$

$\sigma_1 = 1$

$\sigma_1 = \sqrt{2}$

$\rho = \frac{1}{\sqrt{2}}$.

Definition.

The set of points of the plane (x, y) with coordinates given by $x = m_1(y_k), y = y_k$ (or) $x = m_1(y), y$ is called ***the regression curve of the random variable X on the random variable Y***.

The set of points of the plane (x, y) with coordinates given by $x = x_i, y = m_2(x_i)$ $(or) x, y = m_2(x_i)$ is called ***the regression curve of the random variable Y on the random variable X***.

Remark:

1. If all the points of the regression curve lie on a straight line, then there is a linear regression.
2. If X & Y are independent

$$m_2(x) = E(Y|X = x) = E(Y)$$
$$m_1(y) = E(X|Y = y) = E(X)$$

Here,

$m_2(x)$ is independent of x then the regression curve of Y on X lies on a line parallel to the x-axis.

Similarly,

The regression curve of X on Y lies on a line parallel to the y-axis.

These, two lines intersect at the point with coordinates (m_1, m_2).

Example 2.

Find the regression curves for the two-dimensional normal distribution given in example 1.

Solution.

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y)dy = \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\left(\frac{x^2 - 2xy + 2y^2}{2}\right)} dy$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-\left(y-\frac{x}{2}\right)^2} dy$$

$$= \frac{1}{2\pi} e^{-\frac{x^2}{2}} \int_{-\infty}^{\infty} e^{-t^2} dt$$

$$= \frac{1}{2\pi} e^{-\frac{x^2}{2}} 2\int_{-\infty}^{\infty} e^{-t^2} dt$$

$$= \frac{1}{\pi} e^{-\frac{x^2}{2}} \frac{\sqrt{\pi}}{2}$$

$$f_1(x) = \frac{1}{2\sqrt{\pi}} e^{-\frac{x^2}{2}}$$

$$f_2(x) = \int_{-\infty}^{\infty} f(x,y)dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{y^2}{2}} e^{\frac{(x-y)^2}{2}} dx$$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \left(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} dt\right) \qquad t = x - y$$

$$dt = dx$$

$$f_2(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$$

We know that

$$f(y|x) = \frac{f(x,y)}{f_1(x)} \;\; \& \; f(x|y) = \frac{f(x,y)}{f_2(x)}$$

$$f(y|x) = \frac{\frac{1}{2\pi} e^{-\left(\frac{x^2-2xy+2y^2}{2}\right)}}{\frac{1}{\sqrt{2\pi}} e^{-\left(\frac{x^2}{4}\right)}}$$

$$= \frac{\frac{1}{2\pi} e^{-\left(\frac{x^2}{4}\right)} e^{-\left(y-\frac{x}{2}\right)^2}}{\frac{1}{2\sqrt{\pi}} e^{-\left(\frac{x^2}{4}\right)}}$$

$$f(y|x) = \frac{1}{\sqrt{\pi}}\ e^{-\left(y-\frac{x}{2}\right)^2}$$

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

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

$$m_2(x) = \int_{-\infty}^{\infty} y\,\frac{f(x,y)}{f_1(x)}dy = \int_{-\infty}^{\infty} y\ f(y|x)dy$$

$$= \int_{-\infty}^{\infty} y\ \frac{1}{\sqrt{2\pi}}e^{-\left(y-\frac{x}{2}\right)^2}\ dy$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} y\ e^{-\left(y-\frac{x}{2}\right)^2}\ dy \qquad t = y - \frac{x}{2}$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\left(t+\frac{x}{2}\right)e^{-t^2}\ dt \qquad dt = dy$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty} te^{-t^2}\ dt + \frac{1}{\sqrt{\pi}}\frac{x}{2}\int_{-\infty}^{\infty} e^{-t^2}\ dt \qquad y = t + \frac{x}{2}$$

$$= 0 + \frac{1}{\sqrt{\pi}}2\,\frac{x}{2}\int_{0}^{\infty} e^{-t^2}\ dt$$

$$= \frac{1}{\sqrt{\pi}}x.\frac{\sqrt{\pi}}{2}$$

$$m_2(x) = \frac{x}{2}$$

$$m_1(x) = \int_{-\infty}^{\infty} xf\left(\frac{x}{y}\right)dx$$

$$= \int_{-\infty}^{\infty} x\,\frac{1}{\sqrt{2\pi}}e^{-\frac{(x-y)^2}{2}}\ dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} x\ e^{-\frac{(x-y)^2}{2}}dx$$

$$= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(t+y)e^{-\frac{t^2}{2}}\,dt$$

$\therefore m_2(x) = \frac{x}{2}$ & $m_1(y) = y$

The regression curves are straight lines.

Example 4.

The random variable (X, Y) can take on the pairs of values $(x_k, y_l)(k, l = 1,2,3,4,5)$, where $x_1 = 1, x_2 = 2, x_3 = 3, x_4 = 4, x_5 = 5, y_1 = 1, y_2 = 2, y_3 = 3, y_4 = 4, y_5 = 5$. The probabilities P_{kl} for the particular pairs (x_k, y_l) are given in below table.

Probabilities P_{kl}						
y_l	x_k					Marginal distribution of the random variable Y
	1	2	3	4	5	
1	1/12	1/24	0	1/24	1/30	1/5
2	1/24	1/24	1/24	1/24	1/30	1/5
3	1/12	1/24	1/24	0	1/30	1/5
4	1/12	0	1/24	1/24	1/30	1/5
5	1/24	1/24	1/24	1/24	1/30	1/5
Marginal distribution of the random variable Y	1/3	1/6	1/6	1/6	1/6	1

Conditional distribution of Y under the condition $X = x_k$ where $k = 1,2,3,4,5$ are the given in the below table.

x_k						x_k					
y_l	1	2	3	4	5	y_l	1	2	3	4	5
1	¼	¼	0	¼	1/5	1	5/12	5/24	0	5/24	1/6
2	1/8	¼	¼	¼	1/5	2	5/24	5/24	5/24	5/24	1/6
3	¼	¼	¼	0	1/5	3	5/12	5/24	5/24	0	1/6
4	¼	0	¼	¼	1/5	4	5/12	0	5/24	5/24	1/6
5	1/8	¼	¼	¼	1/5	5	5/24	5/24	5/24	5/24	2/6
Total	1	1	1	1	1						

Find the conditional expected value of one random variable under the condition that the second take on a given condition.

i.e., Find $E(Y|X = x_k)$ $k = 1,2,\ldots..,5$

$$E(X|Y = y_l)\ \ l = 1,2,\ldots..,5$$

(or)

Find the regression curve of Y on X and the regression curve of X on Y

$$E(Y|X = 1) = \frac{1}{4}.1 + \frac{1}{8}.2 + \frac{1}{4}.3 + \frac{1}{4}.4 + \frac{1}{8}.5$$

$$= 2\frac{7}{8}$$

$$E(Y|X = 2) = 2\frac{3}{4}$$

$$E(Y|X = 3) = 3\frac{1}{2}$$

$$E(Y|X = 4) = 3$$

$$E(Y|X = 5) = 3$$

Similarly,

$$E(X|Y = 2) = \frac{5}{24}.1 + \frac{5}{24}.2 + \frac{5}{24}.3 + \frac{5}{24}.4 + \frac{1}{6}.1$$

$$= 2.\frac{11}{12}$$

$$E(X|Y = 1) = 2.\frac{1}{2}$$

$$E(X|Y = 3) = 2.\frac{7}{24}$$

$$E(X|Y = 4) = 2.\frac{17}{24}$$

$$E(X|Y = 5) = 2.\frac{11}{12}$$

1. Consists of the points with coordinates $x = x_k, y = E(Y|X = x_k)$ $(k = 1,2, \dots.,5)$

   The points of (1) from the regression curve of Y on X

2. Consists of the points with coordinates $y = y_i,\ x = E(X|Y = y_l)$ $(l = 1,2, \dots,5)$

   The points of (2) from the regression curve of X on Y.

Remark.

1. The regression curve of the random variable Y on the random variable X satisfies the relation $E\{[Y - m_2(X)]^2\} = minimum$
   i.e., the mean quadratic deviation of Y form a function
   $u(X)$ gets its minimum when $u(X)$ gets its minimum when $u(X)$ equals $m_2(X)$ with probability 1.

2. Suppose (X, Y) is a two-dimensional random variable of the cts type with density $f(x, y)$

   $E\{[Y - u(X)]^2\} = \int_{-\infty}^{\infty} f_1(x)[Y - u(X)]^2 f(y|x)dy\} \, dx \ \rightarrow (1)$

   R.H.S of (1) takes its minimal value when $u(x) = m_2(x)$.

3. Let $f(x_1, x_2, \dots \dots, x_n)$ be the density function of the random variable $(X_1, X_2, \dots \dots, X_n)$. Suppose that the conditional moment for $l = 1$ exist.

$$E\left(X_1^l \middle| X_2 = x_2, X_3 = x_3, \dots \dots X_m = x_m\right) = \int_{-\infty}^{\infty} x_1^l \frac{f(x_1, x_2, \dots \dots, x_m)}{g(x_2, x_3, \dots \dots x_m)} dx_1$$

$$m_1(x_2, \dots \dots, x_n) = \frac{\int_{-\infty}^{\infty} x_1 f(x_1, x_2, \dots \dots, x_n) dx_1}{\int_{-\infty}^{\infty} f(x_1, x_2, \dots \dots, x_n) dx_1}$$

Definition.

The set of points of the n-dimensional space $(x_1, x_2, \dots \dots, x_n)$ with the coordinates

$$x_1 = m_1(x_2, x_3, \dots \dots, x_n),\ x_2, x_3, \dots \dots x_n$$

is called the ***regression surface of the 1st type of the random variable X, on the random variables*** $\boldsymbol{X_2, X_3, \dots \dots X_n}$.

UNIT – III

CHARACTEREISTICS FUNCTIONS

3.1. Properties of characteristics functions.

Let X be a random variable and let F(x) be its distributive function.

Definition. The function

$$\phi(t) = E\left(e^{itX}\right)$$

Where t is a real numbers and i is the imaginary unit is called the characteristics function of a random variable X or of the distribution function $F(x)$.

Definition. If X is of random variable of the discrete type with jump points $x_k(k = 1,2,\dots)$ and $P(X = x_k) = p_k$, the characteristic function of X has the form

$$\phi(t) = E\left(e^{itX}\right) = \sum_k p_k e^{itx_k} \dots\dots\dots.(1)$$

Since $\left|e^{itx_k}\right| = 1$ and $\sum_k p_k = 1$, the series on the right hand side of (1) is absolutely and uniform convergent. Thus, the characteristics function $\phi(t)$, as the sum of uniformly convergent series of continuous function, is continuous for every real value of t.

Thus $\phi(t)$ is continuous for every t.

Problem 1. The random variable X can take on the value $x_1 = -1$ and $x_2 = 1$ with probabilities $P(X = -1) = P(X = +1) = 0.5$. Find the characteristic function of this random variable.

Solution.

$\phi(t) = \sum_k p_k e^{itx_k}$

$= P(X = -1)e^{itx_1} + P(X = 1)e^{itx_2}$

$= 0.5\, e^{-it} + 0.5e^{it}$

$= 0.5[\cos t - \sin t + \cos t + \sin t\,]$

$= 0.5(2\cos t)$

$= \cos t$

$i.e., \phi(t) = \cos t$.

Definition. If X is a random variable of the continuous type with density function $f(x)$, its characteristic function is given by

$$\phi(t) = E\left(e^{itX}\right) = \int_{-\infty}^{\infty} f(x)e^{itx}dx \quad \ldots..(1)$$

Since $\int_{-\infty}^{\infty} f(x)\left|e^{itx}\right| dx = \int_{-\infty}^{\infty} f(x)dx = 1.$

∴ the integral in equation (1) is absolutely and uniformly convergent.

Hence $\phi(t)$ is a continuous function for every t.

Problem 2. The density function $f(x)$ defined as

$$f(x) = \begin{cases} 0 & if\ x < 0 \\ 1 & if\ 0 \le x \le 1 \\ 0 & if\ x > 1 \end{cases}$$

This distribution function is called uniform or rectangular. Find its characteristic function.

Solution.

$\phi(t) = \int_{-\infty}^{\infty} f(x)e^{itx}dx$

$= \int_0^1 e^{itx}dx$

$= \left(\frac{e^{itx}}{it}\right)_0^1$

$= \frac{e^{it}-1}{it}$

Properties of characteristic functions

1) $\phi(0) = 1.$

   For,

   $\phi(t) = E(e^{itx})$

   $\phi(0) = E(e^{0}) = E(1) = 1$.

2) $|\phi(t)| \leq 1.$

   For,

   $|\phi(t)| = |E(e^{itX})|$

   $= |\sum_k p_k e^{itX}|$

   $\leq \sum_k p_k |e^{itx_k}|$

   $= E(|e^{itx_k}|)$

   $= E(1)$

   $= 1$

   $\therefore |\phi(t)| \leq 1.$

3) $\phi(-t) = \overline{\phi(t)}.$

   Proof.

   LHS:

   $\phi(-t) = E(e^{itX})$

   $= E(\cos t X - i \sin t\ X)$

   $= E(\cos t\ X) - i\ E(\sin t X)$ ......(1)

   RHS:

   $\overline{\phi(t)} = \overline{E(e^{\imath tX})}$

   $= \overline{E(\cos t X + \imath \sin t X)}$

   $= \overline{E(\cos t\ X) + \imath\ E(\sin t X)}$ ... ... (2)

Using (1) and (2),

$\phi(-t) = \overline{\phi(t)}$

Where $\overline{\phi(t)}$ denotes the complex number conjugate to $\phi(t)$.

Remark:

Every characteristic function satisfies the properties of the characteristic functions. But converse need not be true.

$i.e.$, every function $\phi(t)$ satisfying a properties is need not be a characteristic function of some random variable.

Theorem 3.1. Let the function $\phi(t)$ defined for $-\infty < t < \infty$ such that $\phi(0) = 1$. The function $\phi(t)$ is the characteristic function of some distribution function iff

(i) $\phi(t)$ is continuous.

(ii) For $n = 1,2, \dots$ and every real $t_1, t_2, \dots, t_n$ and complex $a_1, a_2, \dots, a_n$ such that

$$\Sigma_{j,k=1}^{n} \phi(t_j - t_k)\, a_j\, \overline{a_k} \geq 0$$

Proof.

Let $\phi(0) = 1$ and $|e^{itx_k}| = 1$ and $\sum_k p_k = 1$.

$\therefore \phi(t)$ is continuous.

For $n = 1,2, \dots$and every real $t_1, t_2, \dots, t_n$ and complex $a_1, a_2, \dots, a_n$.

Then, $\Sigma_{j,k=1}^{n} \phi(t_j - t_k) a_j\, \overline{a_k} \geq 0$.

$\therefore \phi(t)$ is defined for $-\infty < t < \infty$ and $\phi(0) = 1$.

$\therefore \phi$ is a characteristic function.

Hence proved.

3.2. Characteristic Function and Moments

Theorem 3.2. If the l^{th} moment $m_1$of a random variable exists, it is expressed as $m_l = \frac{\phi^{(l)}(0)}{i^i}$, where $\phi^l(0)$ is the l^{th}derivative of the characteristic function of this random variable at $t = 0$.

Proof.

Consider the random variable X and suppose that its l^{th} moment $m_l = E(X^l)$ exists.

Suppose X is the random variable of the discrete type with jump points x_k.

The characteristic function of X is $\phi(t) = \Sigma_k\, p_k e^{itx_k} \dots (1)$

Differentiate (1) with respect to 't' for l times we get,

$\phi'(t) = \Sigma_k \, p_k e^{itx_k}(it)$

$\phi''(t) = \Sigma_k \, p_k e^{itx_k}(it)^2$

⋮

$\phi^l(t) = \Sigma_k \, p_k e^{itx_k}(it)^l$.

$\therefore \phi^l(t) = E\left(i^l X^l e^{itX}\right) \dots\dots (2) \; (\because E\big(g(x)\big) = \Sigma_k \, p_k g(x_k) \,)$

Since $m_l = E(X^l)$ is exists.

$\Sigma_k \left| p_k i^l x_k^l e^{itX_k} \right| = \Sigma_k \left| p_k x_l^l \right| < \infty$

$\therefore$ RHS of (2) exists.

$\therefore \phi^l(t)$ is exists.

Suppose $f(x)$ is the density function of the random variable X of the continuous type.

$$\phi(t) = E\left(e^{itX}\right) = \backslash limit \int_{\infty}^{\infty} f(x) e^{itx} dx \dots \dots (3)$$

Differentiate (3) $'l'$ times with respect to t we get,

$$\phi^l(t) = \int_{-\infty}^{\infty} f(x) i^l c^l e^{itx} dx$$

$$= E\left(i^l X^l e^{itX}\right)$$

Since $\int_{-\infty}^{\infty} \left| i^l x^l f(x) e^{itx} \right| dx = \int_{-\infty}^{\infty} |x^l f(x)\,| dx = \beta_l$

By assumption, the absolute moment β_l is finite, ϕ_l is exists.

$$\therefore \phi^{(l)}(t) = E\left(i^l X^l e^{itX}\right)$$

Put $t = 0$,

$$\phi^l(0) = E(i^l X^l e^0) = E(i^l X^l) = i^l E(X^l) i^l m_l$$

$\therefore m_l = \frac{\phi^{(l)}(0)}{i^l}$

Hence Proved.

Example 1. Suppose that the random variable X has a Poisson distribution. i.e., it can taken on the values $x_k = k_1$, where k is any non-negative integer, and the probability function is given by the formula

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

where λ is the positive constant. Find the characteristic function of X and moments.

Solution.

$\phi(t) = \Sigma_k p_k e^{itx_k}$

$= \Sigma_{k=0}^{\infty} \left[e^{itk} \frac{\lambda^k}{k!}\right] e^{-\lambda}$

$e^{-\lambda} \Sigma_{k=0}^{\infty} \left[\frac{(\lambda e^{it})^k}{k!}\right]$

$= e^{-\lambda} e^{\lambda} e^{it}$

$\therefore \phi \quad (t) = e^{\lambda(e^{it}-1)}$

$\phi'(t) = e^{-\lambda} e^{\lambda e^{it}} \left(\lambda i e^{it}\right)$

$\therefore \phi'(t) = \lambda i e^{it} e^{\lambda(e^{it}-1)}$

$\phi''(t) = \lambda i e^{it}(i) e^{\lambda(e^{it}-1)} + \lambda i e^{it} \left(\lambda i e^{it} e^{\lambda(e^{it}-1)}\right)$

$= -\lambda e^{it} e^{\lambda(e^{it}-1)} \left(\lambda e^{it} + 1\right)$

We know that , $m_l = \frac{\phi^{(l)}(0)}{(i)^l}$

$\therefore m_1 = \frac{\phi'(0)}{(i)'}$

$= \frac{\lambda i e^{o} e^{\lambda(1-1)}}{i}$

$= \frac{\lambda i}{i} = \lambda$

$i.e., m_1 = \lambda$

$m_2 = \frac{\phi''(0)}{i^2}$

$= \frac{-\lambda e^0 e^{\lambda(1-1)}(\lambda e^0+1)}{-1}$

$= \frac{-\lambda(\lambda+1)}{-1}$

$i.e., m_2 = \lambda(\lambda+1)$

Central moment $\mu_1 = 0$.

The central moment of the second order is

$$\sigma^2 = \mu_2 = m_2 - m_1^2 = \lambda(\lambda+1)^2 - \lambda^2 \ = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Example 2. Find the characteristic function and moment of the normal distribution.

Solution.

Let X be the normal distribution and the density function be $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

$\phi(t) = \int_{-\infty}^{\infty} e^{itx} f(x)dx$

$= \int_{-\infty}^{\infty} e^{itx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{itx} e^{-\frac{x^2}{2}} dx$

$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} e^{-\frac{t^2}{2}} dx$

$= e^{-\frac{t^2}{2}} \left(\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-it)^2}{2}} dx \right)$

$\therefore \phi(t) = e^{-\frac{t^2}{2}}$

$\phi'^{(t)} = -te^{-\frac{t^2}{2}}$

$\phi''^{(t)} = e^{-\frac{t^2}{2}}(t^2 - 1)$

We know that $m_l = \frac{\phi^{(l)}(0)}{(i)^l}$

$\therefore \ m_1 = \frac{\phi'(0)}{(i)^1} = 0$

$m_2 = \frac{\phi''(0)}{(i)^2} = \frac{e^0(0-1)}{-1} = \frac{-1}{-1} = 1$

Clearly all odd order moments are zero and that the even order moments are expressed by the formula

$$m_{2l} = 1.3.5 \ldots (2l-1)$$

Characteristic function of linear transformation

1. If the random variable is translated by a constant n, then characteristic function is multiplied factor e^{itb}. Let $Y = X + b$, where X is random variable and characteristic function is $\phi(t)$.

   Let $\phi_1(t)$ be the characteristic function of Y.

$$\phi_1(t) = E\left(e^{itY}\right)$$
$$= E\left(e^{it(X+b)}\right)$$
$$= E\left(e^{itX}e^{itb}\right)$$
$$= e^{itb}E\left(e^{itX}\right)$$
$$\therefore \phi_1(t) = e^{itb}\phi(t).$$

2. The characteristic function of a random variable aX equals the characteristic function of the random variable X at the point at.

   Let $Y = aX$, where X is a random variable and the characteristic function is $\phi(t)$.

$$\phi_1(t) = E\left(e^{itY}\right) = E\left(e^{itaX}\right) = \phi(at)$$
$$\therefore \phi_1(t) = \phi(at)$$

   In particular, if $a = -1$, we obtain

$$\phi_1(t) = \phi(-t) = \overline{\phi(t)}$$
$$\therefore \phi_1(t) = \overline{\phi(t)}$$

3. Find the characteristic function of the random variable X and Y by $\phi(t)$ and $\phi_1(t)$ respectively, we obtain

$$\phi_1(t) = E\left(e^{itY}\right) = E\left(e^{it(aX+b)}\right) = E\left(e^{itX}e^{itb}\right) = e^{itb}\phi(at)$$
$$\therefore \phi_1(t) = e^{itb}\phi(at)$$

Note:

Suppose $Y = \frac{X-m_1}{\sigma}$

Where m_1 and σ denote the expected value and standard deviation of X.

Here, $a = \frac{1}{\sigma}, b = \frac{-m_1}{\sigma}$

$\phi_1(t) = e^{itb}\phi(at)$

$= e^{\frac{itm_1}{\sigma}}\phi(\frac{t}{\sigma})$

3.3.Semi-Invariants

Definition. Let $\psi(t) = \log\phi(t)$, where $\phi(t)$ is the characteristic function of the random variable under consideration.

Let us expand the function $\phi(t)$ in a power series function in a neighbourhood of $t = 0$,

$$\phi(t) = 1 + \Sigma_{s=1}^{\infty} \frac{m_s}{s!}(it)^s \dots\dots(1)$$

Let z denote the series of RHS of (1),

$$\psi(t) = \log\phi(t) = \log(1 + z)$$

Let us expand the function $\psi(t)$ into a power series,

$$\psi(t) = \frac{z}{1} - \frac{z^2}{2} + \frac{z^3}{3} - \cdots$$

$$= \Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s$$

$$\log\phi(t) = \Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s \dots\dots(2)$$

From (1) and (2),

$$1 + \Sigma_{s=1}^{\infty} \frac{m_s}{s!}(it)^s = e^{\left[\Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s\right]}$$

$$\therefore \phi(t) = 1 + \Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s + \frac{1}{2!}\left[\Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s\right]^2 + \frac{1}{3!}\left[\Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s\right]^3 + \cdots$$

$$\dots\dots.(*)$$

Definition. Let $\psi(t) = 1 + \Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s \dots\dots(1)$

The coefficient k_s in (1) are called ***semi-invariants.***

Result. Derive the semi-invariants in terms of the moments (or) the moments in terms of the semi-invariants.

Proof.

We know that, $\phi(t) = 1 + \Sigma_{s=1}^{\infty} \frac{m_s}{s!}(it)^s \ldots \ldots (1)$

$$i.e., \phi(t) = 1 + \Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s + \frac{1}{2!}\left[\Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s\right]^2 + \frac{1}{3!}\left[\Sigma_{s=1}^{\infty} \frac{k_s}{s!}(it)^s\right]^3 + \cdots (2)$$

Compare the $(it)^s$ for particular values of s in equation (2) we obtain,

$$\left.\begin{aligned} k_1 &= m_1 \\ k_2 &= m_2 + m_1^2 = \sigma^2 \\ k_3 &= m_3 - 3m_1 m_2 + 2m_1^3 \\ k_4 &= m_4 - 3m_2^2 - 4m_1 m_3 + 12m_1^2 m_2 - 6m_1^4 \\ &\ldots \ldots \end{aligned}\right\} \ldots \ldots (3)$$

and

$$\left.\begin{aligned} m_1 &= k_1 \\ m_2 &= k_2 + k_1^2 \\ m_3 &= k_3 + 3k_1 k_2 + 2k_1^3 \\ m_4 &= k_4 + 3k_2^2 + 4k_1 k_3 + 6k_1^2 k_2 + k_1^2 \\ &\ldots \ldots \end{aligned}\right\} \ldots \ldots (4)$$

The semi-invariants can also be in terms of the central moments.

$$k_1 = m_1$$

$$k_2 = \mu_2 = \sigma^2$$

$$k_3 = \mu^3$$

$$k_4 = \mu_4 - 3\mu_2^2$$

$$\ldots \ldots$$

Note.

1. From (3) and (4), if the moments of the l^{th} order exists, all the semi-invariants of order not greater than l also exist.

2. Let $Y = X + b$. Let $\phi(t)$ and $\phi_1(t)$ be the characteristic function of the random variables X and Y respectively, we have

$$\log \phi_1(t) = bit + \log \phi(t)$$

Thus the translation changes only the coefficient of the terms with it to the first power in the expansion (*).Hence it changes only the semi-invariant of the first order.

Example 1. Compute the semi-invariants of the Poisson distribution and moments.

Solution.

The characteristic function of the Poisson distribution is

$$\phi(t) = e^{\lambda(e^{it}-1)}$$

$$\psi(t) = \log \phi(t)$$

$$= \log e^{\lambda(e^{it}-1)}$$

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

$$= \lambda\left(\Sigma_{k=0}^{\infty} \frac{(it)^k}{k!} - 1\right)$$

$$= \lambda\left(1 + \Sigma_{k=1}^{\infty} \frac{(it)^k}{k!} - 1\right)$$

$$i.e., \psi(t) = \lambda \Sigma_{k=1}^{\infty} \frac{(it)^k}{k!}$$

We know that $\psi(t) = \Sigma_{s=1}^{\infty} \frac{k_s}{s!} (it)^s$

$\therefore \mathrm{k}_k = \lambda \ (\mathrm{k} = 1,2, \ldots)$ …..(1)

Using the formulas for the relations between semi-invariants and moments we can obtain from formula (1) the moments of arbitrary order of the Poisson distribution are:

$$m_1 = k_1 = \lambda$$

$$m_2 = k_2 + k_1^2 = \lambda + \lambda^2 = \lambda(1 + \lambda)$$

$$m_3 = k_3 + 3k_1k_2 + k_1^3 = \lambda + 3\lambda^2 + \lambda^3$$

3.4. The Characteristic function of the sum of independent random variables

Let X and Y be two independent random variable. We have e^{itX} and e^{itY} are independent.

Theorem 3.4. The characteristic function of the sum of an arbitrary finite number of independent random variables equals the product of their characteristic functions.

Proof.

We shall find the characteristic function of the sum $Z = X + Y$.

Let $\phi(t), \phi_1(t)$ and $\phi_2(t)$ be denote the characteristic function of the random variables Z, X and Y. We have,

$$\phi(t) = E\left(e^{itZ}\right) = E\left(e^{it(X+Y)}\right) = E\left(e^{itX}.e^{itY}\right)$$

$$= E\left(e^{itX}\right).E\left(e^{itY}\right) = \phi_1(t)\phi_2(t)$$

Corollary. The characteristic function of the sum of n independent random variables equals the product of their characteristic functions.

Proof.

We prove the result the induction on number of random variables.

Let $n = 2$

Let X_1 and X_2 be the two independent random variables, then by the above the theorem we have

$$\phi(t) = \phi_1(t)\phi_1(t)$$

Assume that $X_1, X_2, \dots, X_{n-1}$ are the independent random variables for which characteristic functions $\phi_1(t), \phi_2(t), \dots, \phi_{n-1}(t)$.

Let $Z = X_1 + X_2 + \cdots + X_{n-1}$ and $\phi(t)$ be the characteristic function of Z.

Assume that $\phi(t) = \phi_1(t)\phi_2(t) \dots \phi_n(t)$

Let $Z = X_1 + X_2 + \cdots + X_n$ and $\phi(t), \phi_1(t), \phi_2(t), \dots, \phi_{n-1}(t)$ denote the characteristic function of $Z, X_1, X_2, \dots, X_n$.

$$\phi(t) = E\left(e^{itZ}\right) = E\left(e^{it(X_1+X_2+\cdots+X_n)}\right) = E\left(e^{itX_1}.e^{itX_2}\dots e^{itX_n}\right)$$

$$i.e., \phi(t) = \phi_1(t) + \phi_2(t) + \cdots \phi_n(t).$$

Example 1. Suppose two independent random variables $X_1$and X_2 have Poisson distribution $P(X_1 = r) = \frac{\lambda_1^r}{r!}e^{-\lambda_1}, P(X_2 = r) = \frac{\lambda_2^r}{r!}e^{-\lambda_2}(r = 0,1,2, \dots)$. Consider the random variable $Z = X_1 - X_2$.Determine the characteristic function and semi-invariants of Z.

Solution.

Let $X_1$and X_2 be the two independent random variables having Poisson distribution.

The characteristic function of $X_1$and X_2 are

$$\phi_1(t) = e^{\lambda_1(e^{it}-1)} \text{ and } \phi_2(t) = e^{\lambda_2(e^{it}-1)}$$

The characteristic function of $-X_2$ is

$$\phi_1(-t) = e^{\lambda_2(e^{-it}-1)}$$

Since X_1 and $-X_2$ are independent, we obtain the characteristic function of Z

$\phi(t) = \phi_1(t)\phi_2(-t)$

$= e^{\lambda_1(e^{it}-1)}e^{\lambda_2(e^{-it}-1)}$

$= e^{\lambda_1(e^{it}-1)+\lambda_2(e^{-it}-1)}$

$= e^{\lambda_1\left(1+\frac{(it)}{1!}+\frac{(it)^2}{2!}+\cdots-1\right)+\lambda_2\left(1-\frac{(it)}{1!}+\frac{(it)^2}{2!}-\cdots-1\right)}$

$= e^{\lambda_1\left(\frac{(it)}{1!}+\frac{(it)^2}{2!}+\cdots\right)+\lambda_2\left(-\frac{(it)}{1!}+\frac{(it)^2}{2!}-\cdots\right)}$

$= e^{(\lambda_1-\lambda_2)\frac{(it)}{1!}+(\lambda_1+\lambda_2)\frac{(it)^2}{2!}+(\lambda_1-\lambda_2)\frac{(it)^3}{3!}+\cdots}$

$\therefore \phi(t) = e^{(\lambda_1-\lambda_2)\frac{(it)}{1!}+(\lambda_1+\lambda_2)\frac{(it)^2}{2!}+(\lambda_1-\lambda_2)\frac{(it)^3}{3!}+\cdots}$,

$$\psi(t) = \log \phi(t) = (\lambda_1 - \lambda_2)\frac{(it)}{1!} + (\lambda_1 + \lambda_2)\frac{(it)^2}{2!} + (\lambda_1 - \lambda_2)\frac{(it)^3}{3!} + \cdots$$

All the semi-invariants of odd order of Z equal $\lambda_1 - \lambda_2$ and all the semi-invariants of even order equal $\lambda_1 + \lambda_2$

$$i.e., k_1 = \lambda_1 - \lambda_2, k_3 = \lambda_1 - \lambda_2, \ldots$$

The expected value and the variance of Z are

$$m_1 = k_1 = \lambda_1 - \lambda_2, \sigma^2 = k_2 = \lambda_1 + \lambda_2$$

Note. The converse of the above theorem is not true.

i.e., the characteristic function of the sum of dependent random variables may equal the product of their characteristic functions.

3.5.Determination of the distribution function by the characteristic function

Theorem 3.5.

Let $F(x)$ and $\phi(t)$ denote respectively the distribution function and the characteristic function of the random variable X. If $a+h$ and $a-h (h>0)$ are continuity points of the distribution function $F(x)$,

$$F(a+h) - F(a-h) = \lim_{T\to\infty} \frac{1}{\pi}\int_{-T}^{T} \frac{\sin ht}{t} e^{-ita}\phi(t)dt \,.$$

Proof.

Let X be a random variable of the continuous type with the density function $f(x)$

Let $J = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin ht}{t} e^{-ita}\phi(t)dt \ldots\ldots (1)$

From the definition of the characteristic function we obtain

$$J = \frac{1}{\pi}\int_{-T}^{T} \frac{\sin ht}{t} e^{-ita}\left[\int_{-\infty}^{+\infty} e^{itx} f(x)dx\right]dt$$

$$= \frac{1}{\pi}\int_{-T}^{T} \left[\int_{-\infty}^{+\infty} \frac{\sin ht}{t} e^{-ita} e^{itx} f(x)dx\right]dt$$

$$= \frac{1}{\pi}\int_{-T}^{T} \left[\int_{-\infty}^{+\infty} \frac{\sin ht}{t} e^{it(x-a)} f(x)dx\right]dt$$

We notice that we can interchange the order of integration since the limits of integration with respect to t are finite and the integral is absolutely convergent with respect to x. Thus

$$\int_{-\infty}^{+\infty}\left|\frac{\sin ht}{t}e^{it(x-a)}\right|f(x)dx=\int_{-\infty}^{+\infty}\left|\frac{\sin ht}{t}\right|f(x)dx\leq h\int_{-\infty}^{+\infty}f(x)dx=h$$

We obtain

$J=\frac{1}{\pi}\int_{-\infty}^{\infty}\left[\int_{-T}^{T}\frac{\sin ht}{t}e^{it(x-a)}f(x)dt\right]dx$

$=\frac{1}{\pi}\int_{-\infty}^{\infty}\left[\int_{-T}^{T}\frac{\sin ht}{t}\{\cos(x-a)t+i\sin(x-a\ )t\}f(x)dt\right]dx$

$=\frac{2}{\pi}\int_{-\infty}^{\infty}\left[\int_{0}^{T}\frac{\sin ht}{t}\cos(x-a)t\,f(x)dt\right]dx$

By the formula

$$\sin A\cos B=\frac{1}{2}[\sin(A+B)+\sin(A-B)]$$

And the substitution $A=ht, B=xt-at$, we obtain

$J=\int_{-\infty}^{\infty}\left[\frac{1}{\pi}\int_{0}^{T}\frac{\sin(x-a+h)t}{t}dt-\frac{1}{\pi}\int_{0}^{T}\frac{\sin(x-a-h)t}{t}dt\right]f(x)\ dx$

$=\int_{-\infty}^{\infty}g(x,T)f(x)\ dx \ldots\ldots (2)$

Where $g(x,T)=\frac{1}{\pi}\int_{0}^{T}\frac{\sin(x-a+h)t}{t}dt-\frac{1}{\pi}\int_{0}^{T}\frac{\sin(x-a-h)t}{t}dt$

It is know from mathematical analysis that the integral $\int_{0}^{T}\left(\sin\frac{x}{x}\right)dx$ is bounded for all $T>0$ and converges to $\frac{\pi}{2}$ as $T\to+\infty$. It follows that the expression $|g(x,T)|$ is bounded and

$\lim_{T\to+\infty}\frac{1}{\pi}\int_{0}^{T}\frac{\sin\alpha t}{t}dt=\begin{cases}\frac{1}{2} & for\ \alpha>0\\ -\frac{1}{2} & for\ \alpha<0\end{cases}$

Here the convergence is uniform with respect to α where $|\alpha|=|x-a\pm h|>\delta>0$.

From this fact we obtain

$$\lim_{T\to+\infty} g(x,T) = \begin{cases} 0 \; for \; x < a-h, \\ \frac{1}{2} \; for \; x = a-h, \\ 1 \; for \; a-h < x < a+h, \\ \frac{1}{2} \; for \; x = a+h, \\ 0 \; for \; x > a+h \end{cases}$$

Taking limit on both sides of (2) we obtain

$$\lim_{T\to\infty} J = \int_{-\infty}^{\infty} \lim_{T\to\infty} g(x,T) f(x)\, dx$$

$$= \int_{a-h}^{+ha} f(x)dx$$

$$= F(a+h) - F(a-h) \ldots\ldots (4)$$

From (1) and (4),

$$F(a+h) - F(a-h) = \lim_{T\to\infty} \frac{1}{\pi} \int_{-T}^{T} \frac{\sin ht}{t} e^{ita} \phi(t) dt.$$

Hence proved.

Remark. If the characteristic function $\phi(t)$ is absolutely over the interval $(-\infty, \infty)$, then the corresponding density function $f(x)$ can be determined by $\phi(t)$. In fact, from the absolutely inerrability of the function $\phi(t)$ it follows that the improper integral in Theorem 3.5.1 exists.

Dividing both sides of equation in Theorem 3.5.1 by $2h$, we have

$$\frac{F(x+h) - F(x-h)}{2h} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t) dt\,.$$

Since the RHS of this equation I a continuous function of x, we obtain

$$F'^{(x)} = f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \phi(t) dt$$

From the absolute and uniform convergence of the last integral it follows that the density $F'(x)$ exists and is c continuous function.

Example 1. The characteristic function of the random variable X is given by the formula $\phi(t) = e^{-\frac{t^2}{2}}$. Find the density function of this random variable.

Solution.

We have $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)dt$

$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{-\frac{t^2}{2}}dt$

$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{(t+ix)^2}{2}}e^{\frac{(ix)^2}{2}}dt$

$= \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{(t+ix)^2}{2}}dt$

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

i.e., $f(x) = \frac{1}{\sqrt{2\pi}}e^{\frac{-x^2}{2}}$.

Example. The joint distribution of the random variable $(X, Y)$is given by the density function

$f(x,y) = \begin{cases} \frac{1}{4}\left(1 + xy(x^2 - y^2)\right) & for\ |x| \le 1\ for\ |x| \le 1\ and\ |y| \le 1 \\ 0 & for\ all\ other\ points \end{cases}$

Solution.

$f_1(x) = \int_{-1}^{1} f(x,y)\,dy$

$= \int_{-1}^{1}\frac{1}{4}\left(1 + xy(x^2 - y^2)\right)dy$

$= \frac{1}{4}\left[\int_{-1}^{1} dy + \int_{-1}^{1} x^3y dy + \int_{-1}^{1} xy^3 dy\right]$

$= \frac{1}{4}\left[(y)_{-1}^{1} + \left(\frac{x^3y^2}{2}\right)_{-1}^{1} + \left(\frac{xy^4}{4}\right)_{-1}^{1}\right]$

$= \frac{1}{4}\left[(1+1) + \left(\frac{x^3(1-1)}{2}\right) + \frac{x(1-1)}{4}\right]$

$= \frac{1}{4} \times 2 = \frac{1}{2}$

$i.e., f_1(x) = \frac{1}{2}$

$f_2(x) = \int_{-1}^{1} f(x,y)\,dy$

$= \int_{-1}^{1} \frac{1}{4}(1 + x^3y + xy^3)\, dx$

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

$= \frac{1}{4}\left[\left(1 + \frac{y}{4} + \frac{y^3}{2}\right) - \left(-1 + \frac{y}{4} + \frac{y^3}{2}\right)\right]$

$= \frac{1}{4}(1 + 1) = \frac{1}{2}$

$i.e., f_2(x) = \frac{1}{2}$

Discrete Type.

If the random variable X is of the discrete type, then its probability function obtained from the characteristic function $p_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-itk}\, \phi(t)dt.$

Example. Find the density function of the random variable whose characteristic function is $\phi_1(t) = \begin{cases} 1 - |t| & for\ |t| \le 1 \\ 0 & for\ |t| > 1 \end{cases}$.

Solution.

$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\phi(t)dt$

$= \frac{1}{2\pi}\int_{-1}^{1} e^{-itx}(1 - |t|)dt$

$= \frac{1}{2\pi}\left[\int_{-1}^{0} e^{-itx}(1 + t)dt + \int_{0}^{1} e^{-itx}(1 - t)dt\right] \dots \dots (1)$

Now,

$\int_{-1}^{0} e^{-itx}(1 + t)dt = \left[(1 + t)\frac{e^{-itx}}{-ix}\right]_{-1}^{0} - \int_{-1}^{0} \frac{e^{-itx}}{-ix} dt$

$= \left[\frac{-1}{ix} + 0\right] + \frac{1}{ix}\left(\frac{e^{-itx}}{-ix}\right)_{-1}^{0}$

$$= \frac{-1}{ix} + \frac{1}{ix}\left(-\frac{1}{ix} + \frac{e^{ix}}{ix}\right)$$

$$= \frac{-1}{ix} - \frac{1}{(ix)^2}\left(1 - e^{ix}\right) \dots\dots (2)$$

$$\int_0^1 e^{-itx}(1-t)dt = \left[(1-t)\frac{e^{-itx}}{-ix}\right]_0^1 - \int_0^1 \frac{e^{-itx}}{-ix}dt$$

$$= \left[\frac{e^{-itx}}{-ix}(1-t)\right] - \frac{1}{-ix}\left(\frac{e^{-itx}}{-ix}\right)_0^1$$

$$= \frac{1}{ix} + \frac{1}{(ix)^2}\left(e^{-ix} - 1\right) \dots\dots (3)$$

Substitute equations (2) and (3) in equation (1),

$$f(x) = \frac{1}{2\pi}\left[-\frac{1}{ix} - \frac{1}{(ix)^2}\left(1 - e^{ix}\right) + \frac{1}{ix} + \frac{1}{(ix)^2}\left(e^{-ix} - 1\right)\right]$$

$$= \frac{1}{2\pi}\left[\frac{1}{x^2}\left(1 - e^{ix}\right) - \frac{1}{x^2}\left(e^{-ix} - 1\right)\right]$$

$$= \frac{1}{2\pi x^2}\left[1 - e^{ix} - e^{-ix} + 1\right]$$

$$= \frac{1}{2\pi x^2}\left[2 - (e^{ix} + e^{-ix})\right]$$

$$= \frac{1}{\pi x^2}\left[1 - \frac{(e^{ix} + e^{-ix})}{2}\right]$$

$$= \frac{1}{\pi x^2}[1 - \cos x]$$

$$\therefore f(x) = \frac{1}{\pi x^2}[1 - \cos x]$$

3.6. Characteristic function of multi-dimensional random vectors

Let (X, Y) be a two-dimensional random vectors and let $F(x, y)$ be its distribution function. Let t and a be two arbitrary real numbers. The characteristics function of the random variable (X, Y) or of the distribution function $F(x, y)$ is defined by the formula $\phi(t, u) = E\left[e^{I(tX+uY)}\right]$.

Example 1. The two-dimensional random variable can take on four pairs of values: $(1, 1), (1, -1), (-1, 1)$ and $(-1, -1)$ with the probabilities

$$P(X=1,Y=1)=\frac{1}{3}, P(X=1,Y=-1)=\frac{1}{3},$$

$$P(X=-1,Y=1)=\frac{1}{6}, P(X=-1,Y=-1)=\frac{1}{6}$$

Find the characteristic function.

Solution.

Clearly X and Y are independent.

The characteristic function of the random variable (X,Y) is

$\phi(t,u)=E\left[e^{i(tX+uY)}\right]$

$= e^{i(t+u)}.P_{11}+e^{i(t-u)}.P_{1(-1)}+e^{i(-t+u)}.P_{(-1)1}+e^{i(-t-u)}.P_{(-1)(-1)}$

$= e^{i(t+u)}.\frac{1}{3}+e^{i(t-u)}.\frac{1}{3}+e^{i(-t+u)}.\frac{1}{6}+e^{i(-t-u)}.\frac{1}{6}$

$=\frac{1}{3}e^{it}e^{iu}+\frac{1}{3}e^{it}e^{-iu}+\frac{1}{6}e^{-it}e^{iu}+\frac{1}{6}e^{-it}e^{-iu}$

$=\frac{1}{3}e^{it}\left(e^{iu}+e^{-iu}\right)+\frac{1}{6}e^{-it}\left(e^{iu}+e^{-iu}\right)$

$=\left(e^{iu}+e^{-iu}\right)\left(\frac{1}{3}e^{it}+\frac{1}{6}e^{-it}\right)$

$=(\cos u+isinu+cosu-isinu)\left(\frac{2e^{it}+e^{-it}}{6}\right)$

$=\frac{1}{6}\times 2\cos u\,(2\,cost+2isint+cost-isint)$

$\phi(t)\ \ =\frac{1}{3}\cos u(3\cos t\ \ t)+isint$.

Properties of characteristic functions of multi-dimensional random variables:

1. $\phi(0,0)=E\left(e^{i(oX+0Y)}\right)=1.$
2. $|\phi(t,u)|=\left|E\left(e^{i(tX+tY)}\right)\right|\leq E\left(\left|e^{i(tX+uY)}\right|\right)=1$

   $\therefore |\phi(t,u)|\leq 1$
3. $\phi(-t,-u)=\overline{\phi(t,u)}.$

   For,

$$\phi(-t,-u) = E\left(e^{-i(tX+uY)}\right) = \overline{\phi(t,u)}$$

Note. If all the moments of order k of a multi-dimensional random variable, then the derivatives $\frac{\partial^k(\phi(t,u))}{\partial t^{k-1}\partial u^l}$ for $l = 0,1,2,\dots,k$ exist and can be obtained from the formula $\frac{\partial^k(\phi(t,u))}{\partial t^{k-1}\partial u^l} = i^k E\left(X^{k-l}Y^l e^{i(tX+uY)}\right) \dots \dots (1)$

Remark. The moment $m_{k-l,l}$ is obtained from the formula $m_{k-l,l} = E(X^{k-l}Y^l) = \frac{1}{i^k}\left[\frac{\partial^k(\phi(t,u))}{\partial t^{k-1}\partial u^l}\right]_{t=0,u=0}$.

For the moments of the first and second order we obtain the expressions

$$m_{10} = \frac{1}{i}\left[\frac{\partial\phi(t,u)}{\partial t}\right]_{t=0,u=0}, m_{01} = \frac{1}{i}\left[\frac{\partial\phi(t,u)}{\partial u}\right]_{t=0,u=0}$$

$$m_{20} = \frac{1}{i^2}\left[\frac{\partial^2\phi(t,u)}{\partial t^2}\right]_{t=0,u=0}, m_{11} = \frac{1}{i^2}\left[\frac{\partial^2\phi(t,u)}{\partial t\partial u}\right]_{t=0,u=0}, m_{02} = \frac{1}{i^2}\left[\frac{\partial^2\phi(t,u)}{\partial u^2}\right]_{t=0,u=0}$$

We obtain the characteristic functions of the marginal distributions of the random variables X and Y from the formula $\phi(t,u) = E\left[e^{I(tX+uY)}\right]$ by putting $t = 0$ or $u = 0$ respectively. Thus

$$\phi(t,0) = E\left(e^{itX}\right) = \phi_1(t)$$

$$\phi(0,u) = E\left(e^{itY}\right) = \phi_2(u)$$

i.e., the marginal distribution of X is $\phi_1(t)$ and the marginal distribution of Y is $\phi_2(u)$.

Theorem 3.6. Let $\phi(t)$ be the characteristic function of the random variable (X,Y). If the rectangle $(a-h \le X < a+h, b-g \le Y < b+g)$ is continuity rectangle, then

$(a-h \le P(X < a+h, b-g \le Y < b+g)$

$$= \lim_{T=\to\infty}\frac{1}{\pi^2}\int_{-T}^{T}\int_{-T}^{T}\frac{\sin\ ht}{t}\frac{\sin su}{u}\ e^{-i(at+bu)}\ \phi(t,u)\ dtdu \dots .(1)$$

Thus, if we know $\phi(t,u)$, (1) allows us to determine the probability $P(x_1 \le X < x_2, y \le Y < y_2)$ for an arbitrary continuity rectangle.

Theorem 3.7. Let $F(x,y), F_1(x), F_2(y), \phi(t,u), \phi_1(t)$ and $\phi_2(u)$ denote the distribution functions and the characteristic function of the random variable $(X,Y), X\ and\ Y$ respectively. The random variable X and Y are independent iff the equation $\phi(t,u) = \phi_1(t)\phi_2(u)$ holds for all real *t* and *u*.

Proof.

Suppose that X and Y are independent.

From the theorem , for any real t and u,

$$\phi(\,t,u) = E\left(e^{i(tX+uY)}\right)$$

$$= E\left(e^{ItX}.e^{iuY}\right)$$

$$= E\left(e^{ItX}\right)E\left(e^{iuY}\right)$$

$$= \phi_1(t)\phi_2(u)\ .$$

Conversely, Suppose $\phi(\,t,u) = \phi_1(t)\phi_2(u)$

If the rectangle $(a-h \le X < a+h, b-g \le Y < b+g)$ is a continuity rectangle, then

$$P(a-h \le X < a+h, b-g \le Y < b+g)$$

$$= \lim_{T\to\infty} \frac{1}{\pi^2}\int_{-T}^{T}\int_{-T}^{T} \frac{\sin ht}{t}\frac{\sin gu}{u}\ e^{-i(at+bu)}\ \phi(t,u)dtdu$$

$$= \left(\lim_{T\to\infty} \frac{1}{\pi^2}\int_{-T}^{T} \frac{\sin ht}{t}\ e^{-iat}\ \phi_1(t)dt\right) + \left(\lim_{T\to\infty} \frac{1}{\pi^2}\int_{-T}^{T} \frac{\sin gu}{u}\ e^{-ibu}\ \phi_2(u)du\right)$$

$$= [F(a+h) - F(a-h)][F(b+g) - F(b-g)]$$

We know that, for every arbitrary points a_1 and a_2 we have,

$$F_1(x_2) - F(x_1) = P(x_1 \le X \le x_2)$$

$$P(x_1 \le X < x_2, y_1 \le Y < y_2) = P(x_1 \le X \le x_2)P(y_1 \le Y \le y_2)\ \text{.....(2)}$$

Which is valid for arbitrary continuity rectangle.

From (2) we get,

$$F(x,y) = F_1(x)F_2(Y)$$

∴ X and Y are independent.

Theorem 3.8. (Cramer-Wold Theorem)

The distribution function $F(x, y)$ of a two two-dimensional random variable (X, Y) is uniquely determined by the class of all one-dimensional distribution function of $tX + uY$ where t and u run over all possible real value.

Proof.

Let $Z = tX + uY$ for all real t and u.

Let $\phi_z(v)$ be the characteristic function of Z.

$\phi_z(v) = E\left(e^{iv(tX+uY)}\right)$

$= E\left(e^{i(vtX+vuY)}\right)$

Put $v = 1$ in (1), then

$\phi_z(1) = E\left(e^{i(tX+uY)}\right) = \phi(t, u)$

∴ $\phi_z(1)$ is the characteristic function of the distribution function $F(x, y)$.

According to the theorem 3.6.1, the function $\phi(t, u)$ is uniquely determines $F(x, y)$.

Hence the theorem is proved.

Note. Let us write

$$P(tX + uY < z) = P(X\cos\alpha + Y\sin\alpha < w)$$

Where

$$\cos\alpha = \frac{t}{\sqrt{t^2 + u^2}}, \sin\alpha = \frac{u}{\sqrt{t^2 + u^2}}, \qquad w = \frac{z}{\sqrt{t^2 + u^2}} \ (0 \le \alpha \le 2\pi)$$

The Cramer-Wold theorem can now be formulated in the following way:

The distribution function $F(x, y)$ is uniquely determined by the distribution functions of the projections of (X, Y) on all straight lines passing through the origin.

3.7.Probabillity Generating functions

Let X be a random variable and let $p_k = P(X = k)(k = 0,1, \dots)$, where $\Sigma_k\, p_k = 1$.

Definition.

The function defined by the formula

$$\psi(s) = \sum_k p_k s^k \text{ ,where } -1 \leq s \leq 1 \ldots \ldots (1)$$

is called the ***probability generality function*** of X.

Clearly, $\psi(1) = \sum_k p_k = 1$

Hence, the series of RHS of (1) is absolutely and uniformly convergent in the interval $|s| \leq 1$.

∴ The generating function is continuous.

Example 1.

The random variable X has a binomial distribution

$$p_k = \binom{n}{k} p^k (1-p)^{n-k} \text{ }, (k = 0,1,2, \ldots n).$$

Find the probability generality function.

Solution.

$$\psi(s) = \sum_k p_k s^k$$

$$= \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \, S^k = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k}$$

$$\psi(s) = (ps + q)^n.$$

Example 2.

The random variable X has a Poisson distribution. Find the $\psi(s)$ or the generating function.

Solution.

We know that, $p_k = e^{-\lambda}\frac{\lambda^k}{k!}, (k = 0,1,2, \dots)$

$$\psi(s) = \Sigma_k\, p_k s^k$$

$$= \Sigma_{k=0}^{\infty}\, e^{-\lambda}\frac{\lambda^k}{k} s^k$$

$$= \Sigma_{k=0}^{\infty}\, e^{-\lambda}\frac{(\lambda s)^k}{k!}$$

$$= e^{-\lambda}\,\Sigma_{k=0}^{\infty}\frac{(\lambda s)^k}{k!}$$

$\psi(S) = e^{-\lambda(1-S)}$.

Moments of the random variable *X* determined by generating function

The moments of the random variable X can be determined by the derivatives at the point 1 of the generating function.

Moments of first order and second order are

$\psi'(S) = \Sigma_k k p_k S^{k-1}$

$\psi''(S) = \Sigma_k k(k-1) p_k S^{k-2}$

$m_1 = \psi'(1) = \Sigma_k k p_k = E(X)$

$m_2 = \psi''(1) = \Sigma_k k(k-1) p_k = E(X^2) - E(X)$

$E(X^2) = \psi''(1) + \psi'(1)$.

UNIT – IV

SOME PROBABILITY DISTRIBUTION

4.1.One-Point and Two-Point Distributions

Definition.

The random variable X has a ***one-point distribution*** if there exists a point x_0 such that

$$P(X = x_0) = 1 \quad \dots\dots(1)$$

(1) gives us the probability function.

The distribution function of this probability distribution is given by the formula

$$F(x) = \begin{cases} 0 & for\ x \leq x_0 \\ 1 & for\ x > x_0 \end{cases}$$

The characteristic function of one-point distribution is obtained from the formula

$$\emptyset(t) = e^{itx_0}$$

Theorem 4.1.

The random variable X has a one-point distribution iff the variance of a random variable X equals zero.

Proof.

Let X be a random variable and has a one-point distribution.

i.e., $P(X = x_0) = 1$

The characteristic function is $\emptyset(t) = e^{itx_0}$.

$\emptyset'(t) = e^{itx_0}(ix_0);\quad \emptyset''(t) = e^{itx_0}(ix_0)^2$

$\emptyset'(0) = ix_0 \qquad ; \quad \emptyset''(0) = (ix_0)^2$

$m_1 = \frac{ix_0}{i} = x_0 \quad ; \quad m_2 = \frac{(ix_0)^2}{i^2} = x_0^2$

$m_1 = x_0 \qquad ; \quad m_2 = x_0^2$

$\therefore m_k = x_0^k \quad \forall k.$

$D^2(X) = m_2 - m_1^2 = x_0^2 - x_0^2 = 0$

$\sigma = D^2(X) = 0$

Conversely, let the variance of a random variable X equals zero

i.e., $D^2(X) = 0$

i.e, $E([X - E(X)]^2) = 0 \qquad \rightarrow (1)$

Since expression $\left(X - E(X)\right)^2$ is non-negative, equation (1) is satisfied only if

$P[X - E(X) = 0] = 1$

$\Rightarrow P[X = E(X)] = 1$

i.e., $P[X = x_0] = 1$ where $x_0 = E(X)$

$\therefore X$ has a one-point distribution.

Definition.

The random variable X has a ***two-point distribution*** if there exist two values x_1 and x_2 such that

$$P(X = x_1) = p, P(X = x_2) = 1 - p, (0 < p < 1) \;\ldots\ldots (1)$$

Put $x_1 = 1$ and $x_2 = 0$ in (1), then we have,

$$P(X = 1) = p; \quad P(X = 0) = 1 - p, (0 < p < 1)$$

This distribution is called the ***zero-one distribution***.

The characteristic function of zero-one distribution

Let X be the random variable and has a zero-one distribution. Find characteristic function of X, central moments and γ.

Proof.

Given X has a zero-one distribution

i.e., $P(X = 1) = p, P(X = 0) = 1 - p, (0 < p < 1)$

The characteristic function $\emptyset(t) = E\left(e^{itX}\right)$

$= pe^{itX_1} + (1 - p)e^{itX_0} = pe^{it} + (1 - p)$

$\emptyset(t) = 1 + p\left(e^{it} - 1\right)$

We know that $m_l = \frac{\emptyset^l(0)}{i^l}$

$\emptyset(t) = 1 + p\left(e^{it} - 1\right)$

$\emptyset'(t) = ipe^{it}$

$\emptyset''(t) = (i)^2 pe^{it}$

$\therefore \emptyset^l(t) = i^l pe^{it}$

$\emptyset^l(0) = i^l p$

$m_l = p \quad \forall k$

Variance, $D^2(X) = m_2 - m_1^2 = p - p^2 = p(1 - p)$

$D^2(X) = p(1-p).$

Since $\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$

$= p - 3p^2 + 3p^3$

$= p(1-p)(1-2p)$

Then we obtain,

$\gamma = \frac{\mu_3}{\mu_2^{3/2}} = \frac{p(1-p)(1-2p)}{p^{3/2}(1-p)^{3/2}}$

If $p = 0.5$, then $\gamma = 0$ since here X has a symmetric distribution.

4.2. The Bernoulli Scheme. The Binomial Distribution.

Relation between zero-one distribution and binomial distribution (or) Bernoulli scheme.

Consider n random experiments.

Let the event A be success with probability (or) failure with probability $q = 1 - p$

The results of the 'n' experiments are independent.

From the 'n' random experiments, event A may occurs 'k' times $(k = 1,1,2, \dots \dots n)$.

Let the number of occurrence of A is a random variable X that can take on the values $k = 0,1, \dots \dots, n,$ where the equality $x = k$ means that in 'n' experiments the event A has occurred 'k' times.

$\therefore X$ has the binomial probability function

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

The distribution function of binomial distribution is

$$F(x) = P(X < x) = \sum_{k<x} \binom{n}{k} p^k (1-p)^{n-k}$$

Where the summation extends over all non-negative integers less than x.

Put $n = 1$. Then, event A occurs 'k' times where $k = 0,1$

$\therefore P(X = 1) = p, P(X = 0) = 1 - p$

$\therefore X$ has zero-one distribution.

Claim: For $n \geq 2$, the binomial distribution obtained from the zero-one distribution.

Let $X_r (r = 1,2, \dots\dots, n)$ be independent random variable with the same zero-one distribution.

The probability function of ever X_r has the form

$$P(X_r = 1) = p, P(X_r = 0) = 1 - p$$

Let $X = X_1 + X_2 + \cdots \dots + X_r$

The random variable X can take values $k = 0,1, \dots\dots, n$.

The event $X = k$ occurs iff k of the n random variable X_r take on the value one and $n - k$ of them take on the value zero.

For k, it may happen $\binom{n}{k}$ different ways.

By the independence of the random variable X_r, we get

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$\therefore X$ has the binomial distribution.

Theorem 4.2

Let $X = X_1 + X_2 + \cdots \ldots . + X_n$ where each X_i are zero-one distribution with characteristic function $\emptyset_1(t), \emptyset_2(t), \ldots \ldots, \emptyset_n(t)$. Find the characteristic function of X, moment, central moment and γ of X.

Proof.

Let $\emptyset(t)$ be the characteristic function of X

Given $X_1, X_2, \ldots \ldots, X_n$ be the zero-one distribution.

The characteristic function

$$\emptyset_i(t) = [1 + p(e^{it} - 1)] \text{ where } i = 1 \text{ to } n.$$

Let $X = X_1 + X_2 + \cdots \ldots . + X_n$

$$\emptyset(t) = E(e^{itX}) = E\left(e^{it(X_1 + X_2 + \cdots \ldots + X_n)}\right)$$

[By theorem: *The characteristic function of the sum of an arbitrary finite number of independent random variables equals the product of their characteristic functions*].

i.e., $\emptyset(t) = \emptyset_1(t)\, \emptyset_2(t) \ldots \ldots \emptyset_n(t)$

$$\emptyset(t) = [1 + p(e^{it} - 1)]^n$$

$$m_l = \frac{\emptyset^l(0)}{(i)^l}$$

$\emptyset(t) = [1 + p(e^{it} - 1)]^n$

$\emptyset'(t) = n[1 + p(e^{it} - 1)]^{n-1}(ip)$

$\emptyset'(0) = n[1 + p(0)]^{n-1}(ip) = nip$

$m_1 = \frac{\emptyset'(0)}{(i)^1} = \frac{npi}{i} = np$

$m_1 = np$

$\emptyset''(t) = n(n-1)(ip)^2\left[1+p\left(e^{it}-1\right)\right]^{n-2}\left(e^{it}\right)^2 + n(ip)ie^{it}\left[1+p\left(e^{it}-1\right)\right]^{n-1}$

$$m_2 = \frac{\emptyset''(0)}{(i)^2} = \frac{[np+n(n-1)p^2](i)^2}{(i)^2} = np + n(n-1)p^2$$

$\mu_1 = 0$

$$\mu_2 = m_2 - m_1^2 = np + n(n-1)p^2 - n^2p^2$$

$$= np + n^2p^2 - np^2 - n^2p^2$$

$\mu^2 = np(1-p)$

$\mu_3 = m_3 - 3m_1m_2 + 2m_1^3$

$\mu_3 = np(1-p)(1-2p)$

$$\gamma = \frac{\mu_3}{\mu_2^{3/2}} = \frac{np(1-p)(1-2p)}{[np(1-p)]^{3/2}} = \frac{np(1-p)(1-2p)}{n^{3/2}\,p^{\frac{3}{2}}(1-p)^{3/2}}$$

$$= n^{1-\frac{3}{2}}\,p^{1-\frac{3}{2}}\,(1-p)^{1-\frac{3}{2}}\,(1-2p)$$

$$= n^{-\frac{1}{2}}p^{-\frac{1}{2}}(1-p)^{-\frac{1}{2}}(1-2p)$$

$$\gamma = \frac{1-2p}{\sqrt{np(1-p)}}\,.$$

Theorem 4.3(Addition theorem for the Binomial distribution)

Let X and Y be two independent random variable with binomial distribution. Let $Z = X + Y$. Find the characteristic function of Z.

Proof

Let $\emptyset(t), \emptyset_1(t), \emptyset_2(t)$ be the characteristic function of Z, X and Y respectively.

$$\emptyset_1(t) = \left[1 + p\left(e^{it} - 1\right)\right]^{n_1}$$

$$\emptyset_2(t) = \left[1 + p\left(e^{it} - 1\right)\right]^{n_2}$$

[By theorem: *The characteristic function of sum of an arbitrary finite number of independent random variables equality to the product of their characteristic function*]

i.e., $\emptyset(t) = \emptyset_1(t)\emptyset_2(t)$

$$= \left[1 + p\left(e^{it} - 1\right)\right]^{n_1}.\left[1 + p\left(e^{it} - 1\right)\right]^{n_2}$$

$$\emptyset(t) = \left[1 + p\left(e^{it} - 1\right)\right]^{n_1+n_2}$$

$\therefore Z$ has the binomial distribution with $n = n_1 + n_2$.

Problem. Find the characteristic function moments and central moment of the random variable $Y = \frac{X}{n}$, where X is the random variable and has the binomial distribution.

Solution.

Let $Y = \frac{X}{n}$ where X is random variable and has the binomial distribution.

The random variable Y can take on the values

$$\frac{k}{n} = 0, \frac{1}{n}, \dots \dots \frac{n-1}{n}, 1$$

Since the probability that $Y = \frac{k}{n}$ is equals to the probability that $X = k$

The probability function of Y is

$P\left(Y = \frac{k}{n}\right) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$

The characteristic function $\phi_Y(t)$ of a random variable Y is given by:

$$\varphi_Y(t) = E[e^{itY}] = E[e^{it\frac{X}{n}}].$$

Since X is binomially distributed, the characteristic function of X, denoted $\varphi_X(t)$, is:

$$\varphi_X(t) = E[e^{itX}] = \left((1 - p) + pe^{it}\right)^n$$

For $Y = \frac{X_n}{Y}$, we substitute $\frac{t}{n}$ into the characteristic function of X:

$$\varphi_Y(t) = E[e^{it\frac{X}{n}}] = \varphi_X\left(\frac{t}{n}\right).$$

Thus, the characteristic function of Y is:

$$\varphi_Y(t) = \left((1 - p) + pe^{i\frac{t}{n}}\right)^n.$$

Moments of Y obtained as follows:

The moments of Y can be derived from its expected values. The k-th moment of Y is $E[Y^k]$. Using the fact that $Y = \frac{X}{n}$=n, we can express this as:

$E[Y^k] = E[\left(\frac{X}{n}\right)^k] = \frac{1}{n^k} E[X^k].$

We know the moments of a binomial random variable X, which are related to the parameters n and p.

The first moment is:

$$E[X] = np.$$

The second moment is:

$$E[X^2] = np(1-p) + (np)^2.$$

For higher moments, we can use the binomial expansion and properties of binomial random variables, but the key point is that $E[Y^k]$ will be the corresponding binomial moments scaled by n^{-k}.

In particular,

$m_1 = p,$

$m_2 = \frac{p}{n} + \frac{n-1}{n}\, p^2,$

$\mu_2 = m_2 - m_1^2 = \frac{p}{n} + \frac{n-1}{n}\, p^2 - p^2 = \frac{p(1-p)}{n}.$

5.4. The Polya and Hyper geometric distribution.

Polya distribution.

Consider an urn with 'b' white and 'c' black balls. Let $b + c = N$.

We draw one ball at random and before drawing the next ball we replace the one we have drawn and add s balls of the same colour. Repeat the procedure 'n' times.

Let X be random variable which takes on the values $k(k = 0,1, \ldots\ldots, n)$ if as a result of 'n' drawings. We draw a white balls 'k' times.

We shall find the probability function of X.

The probability of the successive drawing of k white balls is

$$\frac{b(b+s) \ldots\ldots [b + (k-1)S]}{N(N+S) \ldots\ldots [N + (k-1)S]}$$

Similarly, The probability of drawing k white balls in turn and then $n - k$ black balls is

$$\frac{b(b+s)\ldots\ldots[b+(k-1)s]c(c+s)\ldots\ldots[c+(n-k-1)s]}{N(N+S)\ldots\ldots[N+(n-1)S]} \quad \ldots(1)$$

The expression (1) is the probability of drawing k white and $n-k$ black balls in any order. The order of drawing affects only the order of the terms in the numerator of (1).

Since k white and $n-k$ black balls can be drawn in $\binom{n}{k}$ different ways, we have

$$\therefore P(X=ik)$$

$$=\binom{n}{k}\frac{b(b+s)\ldots\ldots[b+(k-1)s]c(c+s)\ldots\ldots[c+(n-k-1)s]}{N(N+s)\ldots\ldots[N+(n-1)s]}$$

Definition.

The random variable X with the probability distribution

$$P(X=k)=\binom{n}{k}\frac{b(b+s)\ldots\ldots[b+(k-1)s]c(c+s)\ldots\ldots[c+(n-k-1)s]}{N(N+S)\ldots\ldots[N+(n-1)S]} \quad \ldots(2)$$

has a ***Polya distribution***.

Denote $Np=b, Nq=c, N\alpha=s$, where p and q are probabilities of drawing a white and a black ball respectively, on the 1st drawing.

Equation (2) ⇒

$$P(X=k)$$

$$=\binom{n}{k}\frac{p(p+\alpha)\ldots\ldots[p+(k-1)\alpha]q(q+\alpha)\ldots\ldots[q+(n-k-1)\alpha]}{1(1+\alpha)\ldots\ldots[1+(n-1)\alpha]}$$

Clearly, $\sum_{k=0}^{n} P(X=K)=1$

i.e., $\sum_{k=0}^{n}\binom{n}{k}\frac{p(p+\alpha)\ldots\ldots[p+(k-1)\alpha]q(q+\alpha)\ldots\ldots[q+(n-k-1)\alpha]}{1(1+\alpha)\ldots\ldots[1+(n-1)\alpha]}=1$

Find 1st and 2nd moments of X.

$m_1 = E(X)$

$\sum_{k=0}^{n} kP(X = k)$

$$= \sum_{k=0}^{n} k \binom{n}{k} \frac{p(p+\alpha) \ldots \ldots [p+(k-1)\alpha]q(q+\alpha) \ldots \ldots [q+(n-k-1)\alpha]}{1(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$= \sum_{k=0}^{n} k \frac{n(n-1)!}{k(k-1)!\,(n-k)!} \frac{p(p+\alpha) \ldots \ldots [p+(k-1)\alpha]q(q+\alpha) \ldots \ldots [q+(n-k-1)\alpha]}{1(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$= pn \sum_{k=0}^{n} \binom{n-1}{k-1} \frac{(p+\alpha) \ldots \ldots [p+(k-1)\alpha]q(q+\alpha) \ldots \ldots [q+(n-k-1)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

Put $l = k - 1$

$E(X)$

$$= pn \sum_{l=0}^{n-1} \binom{n-1}{l} \frac{(p+\alpha) \ldots \ldots [p+l\alpha]q(q+\alpha) \ldots \ldots [q+(n-l-2)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$\therefore E(X) = pn$

$E(X^2) = \sum_{k=0}^{n} K^2 P(X = k)$

$$= \sum_{k=0}^{n} k^2 \frac{n(n-1)!}{k(k-1)!\,(k-r)!} \frac{p(p+\alpha) \ldots \ldots [p+(k-1)\alpha]q(q+\alpha) \ldots \ldots [q+(n-k-1)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$= np \sum_{k=1}^{n} k \binom{n-1}{k-1} \frac{(p+\alpha) \ldots \ldots [p+(k-1)\alpha]q(q+\alpha) \ldots \ldots [q+(n-k-1)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

Put $l = k - 1$

$$= np \sum_{l=0}^{n-1} (l+1) \binom{n-1}{l} \frac{(p+\alpha) \ldots \ldots [p+l\alpha]q(q+\alpha) \ldots \ldots [q+(n-l-2)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$= np\{\sum_{l=0}^{n-1} l \binom{n-1}{l} \frac{(p+\alpha) \ldots \ldots [p+l\alpha]q(q+\alpha) \ldots \ldots [q+(n-l-2)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$+ \sum_{l=0}^{n-1} \binom{n-1}{l} \frac{(p+\alpha) \ldots \ldots [p+l\alpha]q(q+\alpha) \ldots \ldots [q+(n-l-2)\alpha]}{(1+\alpha) \ldots \ldots [1+(n-1)\alpha]}$$

$$= np(A + B)$$

$$A = \sum_{l=0}^{n-1} \binom{n-1}{l} \frac{(p+\alpha)(P+2\alpha)\ldots\ldots[p+l\alpha]q\ldots\ldots[q+n-l-2]\alpha}{(1+\alpha)\ldots\ldots[1+(n-1)\alpha]}$$

$$= \sum_{l=0}^{n-1} \frac{l(n-1)(n-2)!}{l(l-1)!\,(n-l+1)!} \frac{(p+\alpha)(p+2\alpha)\ldots\ldots(p+l\alpha)q\ldots\ldots[q+n-l-2]\alpha}{(1+\alpha)\ldots\ldots[1+(n-1)\alpha]}$$

$$= \frac{(p+\alpha)(n-1)}{1+\alpha} \sum_{l=1}^{n-1} \binom{n-2}{l-1} \frac{(p+2\alpha)\ldots\ldots(p+l\alpha)q\ldots\ldots[q+n-l-2]\alpha}{[1+(n-1)\alpha]}$$

$$[\because r = l-1, l = r+1, -l = -r-1]$$

$$A = \frac{(P+\alpha)(n-1)}{1+\alpha} \sum_{l=1}^{n-1} \binom{n-2}{l-1} \frac{(p+2\alpha)\ldots\ldots[p+(r+1)\alpha]q\ldots\ldots[q+(n-r-3)\alpha]}{(1+2\alpha)\ldots\ldots[1+(n-1)\alpha]}$$

$$A = \frac{(p+\alpha)(n-1)}{1+\alpha} X$$

Clearly,

$$E(X^2) = np\left[\frac{(p+\alpha)(n-1)}{1+\alpha} + 1\right] = np\left[\frac{np - p + n\alpha - \alpha + 1 + \alpha}{1+\alpha}\right]$$

$$= np\left[\frac{np+n\alpha+1-p}{1+\alpha}\right]$$

$$E(X^2) = np\left[\frac{np+n\alpha+q}{1+\alpha}\right]$$

$$\mu_1 = 0$$

$$\mu_2 = D^2(X) = m_2 - m_2^1$$

$$= np\frac{np+q+n\alpha}{1+\alpha} - n^2p^2 = np\left[\frac{np+q+n\alpha-np(1+\alpha)}{1+\alpha}\right]$$

$$= np\left[\frac{np+q+n\alpha-np-np\alpha}{1+\alpha}\right] = np\left[\frac{q+n\alpha(1-P)}{1+\alpha}\right]$$

$$= np\left[\frac{q+n\alpha q}{1+\alpha}\right]$$

$$D^2(X) = npq\left(\frac{1+n\alpha}{1+\alpha}\right).$$

Remark.

In the Polya scheme ‘s’ may also be negative. Since inequalities

$b + (k-1)S \geq 1$ and $c + (n-k-1)S \geq 1$

must hold, k must then satisfy the double inequality.

$\max\left(0, n - 1 + \frac{c-1}{S}\right) \leq k \leq \min\left(n, \frac{1-b}{S} + 1\right).$

Theorem 4.4

If for $N = 1,2, \dots..$ equality $p = \frac{b}{N} =$ constant is satisfied and $\lim_{N\to\infty} \alpha = 0.$ Then the probability fucniton of the random variable X with Polya distribution tends to the probability function at the binomial distribution.

Proof.

Let N, b and c tends to infinity so that

$$p = \frac{b}{N} = \text{constant}$$

Clearly, $q = 1 - p =$ constant

Suppose that, $\lim_{N\to\infty} \alpha = 0$

We know that $N\alpha = S$

$\Rightarrow \alpha = \frac{S}{N} \Rightarrow \mathit{lim}_{N\to\infty}\, \alpha = \mathit{lim}_{N\to\infty} \frac{S}{N} \Rightarrow 0 = \mathit{lim}_{N\to\infty} \frac{S}{N}$

$\Rightarrow S$ is constant

$P(X = k) = \binom{n}{k} \frac{p(p+\alpha)\dots\dots[p+(k-1)\alpha]q(q+\alpha)\dots\dots[q+(n-k-1)\alpha]}{1(1+\alpha)\dots\dots[1+(n-1)\alpha]}$

$\lim_{N\to\infty} P(X = k) = \binom{n}{k} p^k q^{n-k}$

This is the probability distribution of binomial distribution.

Hyper geometric distribution

Hyper geometric distribution is obtained from Polya distribution by putting $s = -1$,s

$$P(X = k) = \binom{n}{k} \frac{b(b+s)......[b+(k-1)s]c(c+s)......[c+(n-k-1)s]}{N(N+S)......[N+(n-1)S]}$$

Put $Np = n; Nq = c; N\alpha = s = -1$

$$P(X = k) = \binom{n}{k} \frac{NP(NP-1)......(NP-K+1)Nq......(Nq-n+k-1)}{N(N-1).......(N-n+1)}$$

$$P(X = k) = \frac{\binom{Np}{k}\binom{Nq}{n-k}}{\binom{N}{n}}$$

This is the ***probability function of Hyper geometric distribution***.

$m_1 = E(X) = np$

We know that $D^2(X) = npq\frac{1+n\alpha}{1+\alpha}$

$N\alpha = S \Rightarrow N\alpha = 1$

$\Rightarrow \alpha = -\frac{1}{N}$

$$D^2(X) = npq \frac{1-\frac{n}{N}}{1-\frac{1}{N}}$$

$$D^2(X) = npq\frac{N-n}{N-1}.$$

4.5. The Poisson distribution

Definition.

A random variable X with probability function is

$$P(X = r) = \frac{\lambda^r}{r!}\, e^{-\lambda}, r = 0,1,2, \ldots$$

where λ is a positive constant is called the ***Poisson distribution***.

Derive characteristic function and moments and central moments of Poisson distribution

The characteristic function $\varphi_X(t)$ of a random variable X is defined as:

$$\varphi_X(t) = E[e^{itX}].$$

For the Poisson distribution, we can calculate this as follows. Using the definition of the expected value and the PMF of the Poisson distribution:

$$\varphi_X(t) = E\left[e^{itX}\right] = \sum_{k=0}^{\infty} e^{itk} P(X = k) = \sum_{k=0}^{\infty} e^{itk} \frac{\lambda^k e^{-\lambda}}{k!}.$$

We can factor out $e^{-\lambda}$ since it does not depend on *k*:

$$\varphi_X(t) = E\left[e^{itX}\right] = e^{-\lambda} \textstyle\sum_{k=0}^{\infty} \frac{\left(\lambda e^{it}\right)^k}{k!}.$$

The sum is now the Taylor series expansion of $e^{\lambda eit}$:

$$\varphi_X(t) = e^{-\lambda} e^{\lambda e^{it}} = e^{\lambda\left(e^{it}-1\right)}.$$

Thus, the characteristic function of a Poisson random variable X with parameter λ is:

$$\varphi_X(t) = e^{\lambda\left(e^{it}-1\right)}.$$

$\emptyset(t) = e^{\lambda\left(e^{it}-1\right)}$

Then we obtain the moments as

$m_1 = \lambda;\ m_2 = \lambda(\lambda + 1);\ \mu_2 = \lambda$

Theorem 4.5.

Let the random variable X_n have a binomial distribution defined by formula

$$P(X_n = r) = \frac{n!}{r!(n-r)!}\ P^r(1 - P)^{n-r}$$

Where r takes on the values $0,1,2, \ldots\ldots, n$. If for $n = 1,2, \ldots..$ the relation $p = \frac{\lambda}{n}$ holds, where $\lambda > 0$ is a constant, then

$$\lim_{n\to\infty} P(X_n = r) = \frac{\lambda^r}{r!}\ e^{-\lambda}$$

Proof.

$$P(X_n = r) = \frac{n!}{r!(n-r)!}\ (1 - P)^{n-r}$$

Put $P = \frac{\lambda}{n}$

$$P(\lambda_n = r) = \frac{n!}{r!(n-r)!}\left(\frac{\lambda}{n}\right)^r\left(1 - \frac{\lambda}{n}\right)^{n-r}$$

$$= \frac{n(n-1)\ldots\ldots(n-r+1)}{r!}\frac{\lambda^r}{n^r}\left(1 - \frac{\lambda}{n}\right)^n\left(1 - \frac{\lambda}{n}\right)^{-r}$$

$$= \frac{\lambda^r}{r!}\left(1 - \frac{\lambda}{n}\right)^n \frac{n(n-1)\ldots\ldots(n-r+1)}{r!}\frac{1}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$P(\lambda_n = r) = \frac{\lambda^r}{r!}\left(1 - \frac{\lambda}{n}\right)^n \frac{1.\left(1-\frac{1}{n}\right)\ldots..\left(1-\frac{r-1}{n}\right)}{\left(1-\frac{\lambda}{n}\right)^r}$$

$$\lim_{n\to\infty} P(X_n = r) = \frac{\lambda^r}{r!}\lim_{n\to\infty}\left(1 - \frac{\lambda}{n}\right)^n\left(\frac{1.\left(1-\frac{1}{n}\right)\ldots..\left(1-\frac{r-1}{n}\right)}{\left(1-\frac{\lambda}{n}\right)^r}\right)$$

$$\therefore \lim_{n\to\infty} P(X_n = r) = \frac{\lambda^r}{r!}\ e^{-\lambda}$$

Since, $\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda} + \lim_{n\to\infty}\left(\frac{1.\left(1-\frac{\lambda}{n}\right)\ldots\ldots\left(1-\frac{r-1}{n}\right)}{\left(1-\frac{\lambda}{n}\right)^r}\right)$

$\therefore \lim_{n\to\infty} P(X_n = r) = \frac{\lambda^r}{r!}\, e^{-\lambda}$

Hence the proof.

Remark.

In figure 1, there are two graphs, one of binomial distribution with $n = 0.5$ and $p = 0.3, \lambda = 1.5$ and one of the poisson distribution with same $\lambda = 1.5$. In figure 2 represents two such graphs for $n = 10$ and $p = 0.15$, then $\lambda = 1.5$.

$\therefore$ For larger values of n, the binomial and Poisson distribution will almost coincide.

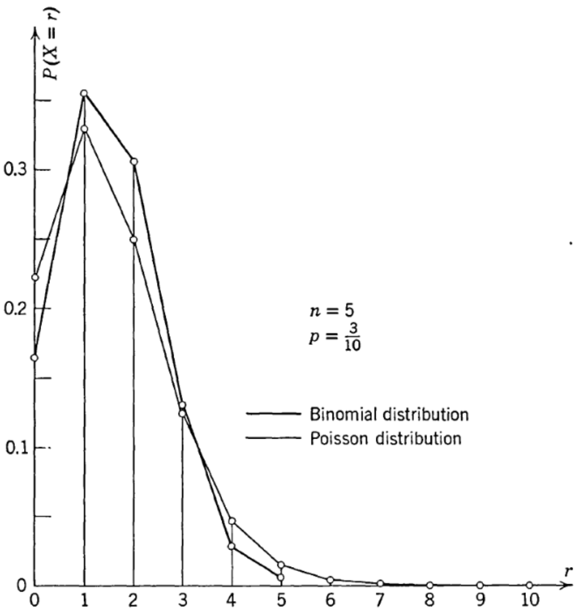


Figure 1

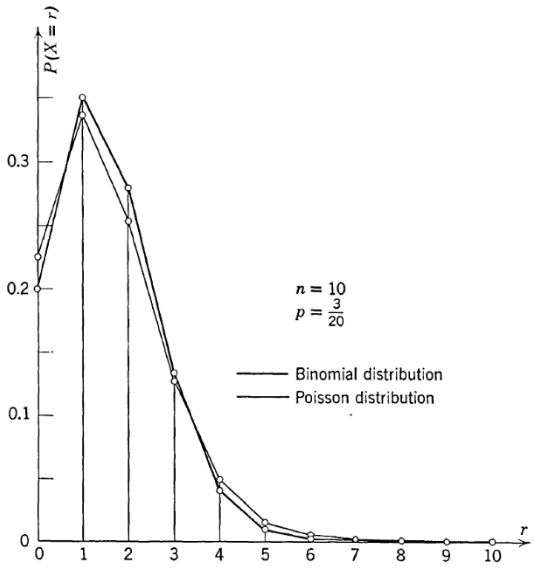


Figure 2

Addition theorem for independent random variable with Poisson distribution

Let the independent random variable X_1 and X_2 have Poisson distribution. Let $X = X_1 + X_2$. Then X has a Poisson distribution.

Proof.

Let the independent random variable X_1 and X_2 have Poisson distribution.

$$P(X_1 = r) = \frac{\lambda_1^r}{r!}e^{-\lambda} ;\ P(X_2 = r) = \frac{\lambda_2^r}{r!}e^{-\lambda} , (r = 0,1,2, \ldots.)$$

The characteristic function of X_1 and X_2 are

$$\emptyset_1(t) = e^{[\lambda_1(e^{it}-1)]} \text{ and } \emptyset_2(t) = e^{[\lambda_2(e^{it}-1)]}$$

Let $X = X_1 + X_2$.

Let $\emptyset(t)$ be the characteristic $\emptyset$ function of X.

Since X_1 and X_2 are independent random variable.

$\therefore \emptyset(t) = \emptyset_1(t)\emptyset_2(t) = e^{[\lambda_1(e^{it}-1)]}e^{[\lambda_2(e^{it}-1)]}$

$i.e., \emptyset(t) = e^{(\lambda_1+\lambda_2)(e^{it}-1)}$

which is a ***characteristic function of the random variable with Poisson distribution*** having the expected value $\lambda_1 + \lambda_2$.

Hence the proof.

4.6. The Uniform Distribution

Definition.

The random variable X has a ***uniform or rectangular distribution*** if its density function $f(x)$ is given by the formula

$$f(x) = \begin{cases} \frac{1}{2h} & for\ a - h \le x \le a + h,\ where\ a\ and\ h \ge 0\ are\ constants. \\ 0 & otherwise \end{cases}$$

The distribution function $F(x)$ of this random variable is given by the formula

$$F(x) = \begin{cases} 0 & for\ \ x < a - h \\ \frac{1}{2}\int_{a-h}^{x} dx = \frac{x-(a-h)}{2h} & for\ a - h \le x \le a + h \\ 1 & for\ x > a + h \end{cases}$$

The characteristic function of X is

$$\emptyset(t) = \int e^{itX} dx$$

$$= \frac{1}{2h}\int_{a-h}^{a+h} e^{itx} dx = \frac{1}{2h}\left(\frac{e^{itx}}{it}\right)_{a-h}^{a+h} = \frac{1}{2h}\frac{e^{it(a+h)} - e^{it(a-h)}}{it}$$

$$= \frac{1}{2h}\frac{e^{ita}\left[e^{ith} - e^{-ith}\right]}{it} = \frac{1}{2h}\frac{e^{ita}[\cos th + isinth - costh + isinth]}{it}$$

$$= \frac{e^{ita}}{2h}\frac{2isinth}{it}$$

$$\emptyset(t) = \frac{e^{ita} sinth}{th}.$$

Moment:

$$m_k = E(X^k)$$

$$= \int x^k f(x) dx$$

$$= \frac{1}{2h}\int_{a-h}^{a+h} x^k dx \text{ s}$$

$$= \frac{1}{2h}\left(\frac{x^{K+1}}{k+1}\right)_{a-h}^{a+h}$$

$$= \frac{1}{2h}\frac{(a+h)^{k+1} - (a-h)^{K+1}}{k+1}$$

$m_1 = \frac{1}{2h}\frac{(a+h)^2-(a-h)^2}{2}$

$= \frac{1}{2h}\frac{a^2+h^2+2ha-a^2-h^2+2ah}{2}$

$= \frac{1}{2h}\frac{4ah}{2} = a$

$i.e., m_1 = a$

$m_2 = \frac{1}{2h}\frac{(a+h)^3-(a-h)^3}{3}$

$= \frac{1}{6h}[a^3 + h^3 + 3a^{2h} + 3ah^2 - (a^3 - h^3 - 3a^2h + 3ah^2)]$

$= \frac{1}{6h}[a^3 + h^3 + 3a^2h + 3ah^2 - a^3 + h^3 + 3a^2h - 3ah^2]$

$= \frac{1}{6h}[2h^3 + 6a^2h] = \frac{1}{6h} \times 2h[h^2 + 3a^2]$

$i.e., \; m_2 = \frac{1}{3}(h^2 + 3a^2)$

$\mu_1 = 0$

$\mu_2 = m_2 - m_1^2 = \frac{1}{3}(h^2 + 3a^2) - a^2 = \frac{1}{3}h^2$

$i.e., \; \mu_2 = \frac{1}{3}h^2.$

4.7. The Normal Distribution

Definition.

The random variable X has a ***normal distribution*** if its density function is

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}}$, where $\sigma > 0$.

Problem 1.

Prove that $f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-m)^2}{2\sigma^2}}$ is a density function and derive characteristic function.

Proof.

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

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-(x-m)^2}{2\sigma^2}} dx$$

Put $Y = \frac{X-m}{\sigma}$

$$dy = \frac{dx}{\sigma}$$

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{\infty}^{\infty} e^{\frac{-(x-m)^2}{2\sigma^2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy$$

$$= 1$$

$$\therefore \int_{-\infty}^{\infty} f(x)dx = 1$$

Hence, $f(x)$ is a density function.

Let the characteristic function of Y be $\emptyset(t)$

$$\emptyset(t) = e^{-\frac{t^2}{2}}$$

Then, the characteristic function of X is

$$\emptyset_1(t) = e^{im - \frac{1}{2}\sigma^2 t^2}$$

Clearly,

$$m_1 = m; m_2 = \sigma^2 + m^2$$

$$\mu_2 = \sigma^2$$

where m is the expected value of X and σ is the standard deviation.

Remark.

1. The shape of the curve of the density of the normal distribution depends on the parameter σ; this curve is called normal curve. It is illustrated in figure, representing three normal distributions with the same expected value $m = 0$ and different standard deviations: $\sigma = 1, \sigma = 0.5$ and $\sigma = 0.25$.
2. The normal distribution with expected value m and standard deviation σ is denoted by $N(m, \sigma)$.
3. By the symmetry of the normal curve with respect to the expected value m all the central moments of odd order vanish

$$\mu_{2k+1} = 0 \quad \forall k.$$

4. $\mu_{2k} = 1.3 \dots \dots (2k-1)\sigma^{2k} \quad (we\ already\ proved)$
5. $P(|X-m| > \lambda_\sigma) = \frac{2}{\sqrt{2\pi}} \int_\lambda^\infty e^{-\frac{y^2}{2}} dy$

$$P(|X-m| > \lambda_\sigma) = P\left(\frac{|X-m|}{\sigma} > \lambda\right) = P(|Y| > \lambda)$$

where $Y = \frac{X-m}{\sigma}$

$$P(X > m + \lambda\sigma) = P(Y > \lambda) = \frac{1}{\sqrt{2\pi}} \int_\lambda^\infty e^{-\frac{y^2}{2}} dy$$

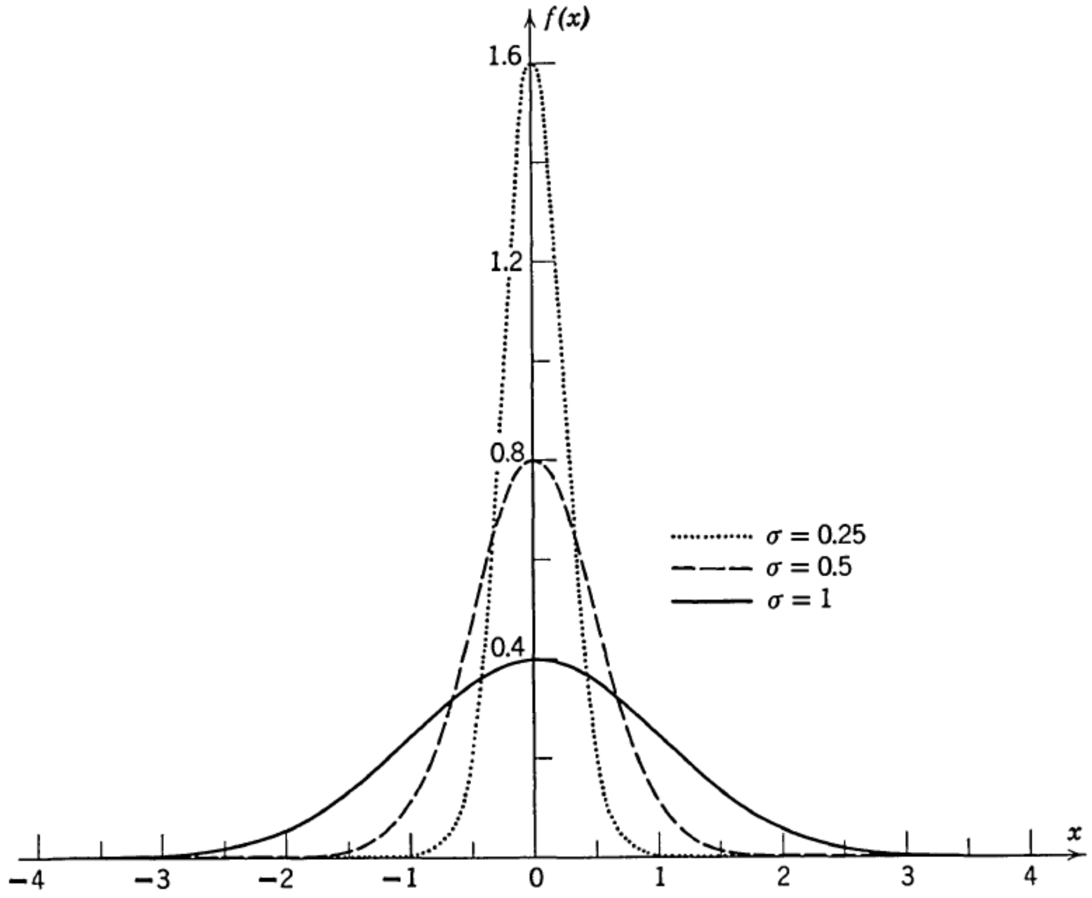


Figure 4.7.1

Problem 2.

The random variable X has the distribution $N(1,2)$. Find the probability that x is greater than 3 in absolute value.

Solution.

To find $P(|x| > 3)$

We have the random variable $Y = \frac{X-m}{\sigma}$

Since $X \sim N(1,2)$, the random variable becomes $Y = \frac{X-1}{2}$

Then

$P(|X| > 3\ ) = P(|2Y + 1| > 3)$

$$= P\left(\left|Y+\frac{1}{2}\right| > \frac{3}{2}\right) \qquad \left(\because \begin{array}{c} Y = \frac{x-1}{2} \\ 2Y+1 = x \end{array}\right)$$

$$\left(\because \left|Y+\frac{1}{2}\right| = \{-\left(Y+\frac{1}{2}\right) + \left(Y+\frac{1}{2}\right)\right)$$

$$= P\left(-\left(Y+\frac{1}{2}\right) > \frac{3}{2}\right) + P\left(\left(Y+\frac{1}{2}\right) > \frac{3}{2}\right)$$

$$= P\left(Y+\frac{1}{2} < -\frac{3}{2}\right) + P\left(Y+\frac{1}{2} > \frac{3}{2}\right)$$

$$= P\left(Y < -\frac{3}{2} - \frac{1}{2}\right) + P\left(Y > \frac{3}{2} - \frac{1}{2}\right)$$

i.e., $P(|X| > 3) = P(Y < -2) + P(Y > 1) \rightarrow (1)$

By the definition of Y, we have

$P(Y > 1) = \frac{1}{\sqrt{2\pi}}\int_1^\infty e^{-\frac{t^2}{2}}dt \cong 0.159 \rightarrow (2) \quad (from\ normal\ table)$

$P(Y < -2) = 1 - P(Y > 2)$

$$= 1 - \frac{1}{\sqrt{2\pi}}\int_2^\infty e^{-\frac{t^2}{2}}dt$$

$$= 1 - 0.977250 \ (from\ normal\ table)$$

$$= 0.02275$$

$P(Y < 2) \simeq 0.023 \rightarrow (3)$

From (2) & (3)

(1)$\Rightarrow P(|X| > 3) = 0.023 + 0.159$

$P(|X| > 3) = 0.182.$

Remark.

1. From the normal table for the above problem

   $P(|X - m| > \sigma \cong 0.3173$

$P(|X - m| > 2\sigma \cong 0.0455$

$P(|X - m| > 3\sigma) \cong 0.0027$

The normal distribution is highly concentration around its expected value. The probability the value of X differs from the expected value by more than 3σ is smaller than 0.01.

This is called 3-sigma rule (or) three-sigma rule.

Theorem 4.7. (Addition theorem for normal distribution).

If X and Y are two independent random variables and X$\sim N(m,\sigma)$ and $Y\sim N(m_2,\sigma_2)$. Then $Z = X + Y$ also has a normal distribution.

Proof.

Given, $X\sim N(m,\sigma_1)$ and $Y\sim N(m,\sigma_2)$

The characteristic function of X and Y are

$$\emptyset_1(t) = e^{m_1 it - \frac{1}{2}t^2\sigma_1^2}, \emptyset_2(t) = e^{m_2 it - \frac{1}{2}t^2\sigma_2^2}$$

Let $\emptyset(t)$ be the characteristic function of Z.Then

$$\emptyset(t) = \emptyset_1(t)\emptyset_2(t) \qquad (\because X_1 \& X_2 \text{ are independent})$$

$$= e^{m_1 it - \frac{1}{2}t^2\sigma_1^2} . e^{m_2 it - \frac{1}{2}t^2\sigma_2^2}$$

$$i.e., \emptyset(t) = e^{(m_1+m_2)it - \frac{1}{2}t^2(\sigma_1^2+\sigma_2^2)}$$

This the characteristic function of the normal distribution $N\left(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right)$

$\therefore X\sim N\left(m_1 + m_2, \sqrt{\sigma_1^2 + \sigma_2^2}\right)$

4.8. The Gamma Distribution

Gamma distribution defined for $p > 0$

$$\Gamma(p) = \int_0^\infty x^{P-1} e^{-x}\, dx \ \ldots (1)$$

(1) is uniformly converges with respect to p and $\Gamma(p)$ is continuous function. Integrating (1) by parts, we obtain

$$\Gamma(p+1) = \int_0^\infty x^P e^{-x}\, dx$$

$$= (-x^p e^{-x})_0^\infty + \int_0^\infty p e^{-x} x^{P-1}\, dx$$

$$= 0 + p\Gamma p$$

$\Gamma(p+1) = p\Gamma p$ …(2)

If $p = n$, where n is an integer, we obtain from (2)

$$\Gamma(n+1) = n\Gamma n$$

$$\Gamma(n) = (n-1)\Gamma(n-1)$$

.

.

$$\Gamma(2) = 1\Gamma(1)$$

Since $\Gamma(1) = \int_0^\infty e^{-x}\, dx = -[e^{-x}]_0^\infty = 1$

$$\therefore \Gamma(n) = n(n-1) \ldots \ldots .1.$$

$$\Gamma(n+1) = n!$$

Remark.

Substitute $y = \frac{x}{a}, (a > 0)\ \ in\ (1)$

$$\Gamma(p) = \int_0^\infty (ya)^{p-1}\, e^{-ay} (ady)$$

$$= a^P \int_0^\infty y^{p-1}\, e^{-ay} dy$$

$$\frac{\Gamma(p)}{a^p} = \int_e^\infty y^{p-1}\, e^{-ay} dy \qquad \rightarrow (3)$$

Equation (2) is also valid when a is complex number $a = b + ic$ where $b > 0$.

Definition.

The random variable X has a ***gamma distribution*** if its density function is

$$f(x) = \begin{cases} 0 & for\ \ x \le 0 \\ \dfrac{b^p}{\Gamma(p)}\, x^{P-i} e^{-bx} & for\ x > 0 \end{cases} \qquad \rightarrow (1)$$

where $b > 0\ and\ p > 0$

$$\int_\infty^\infty f(x) dx = \int_0^\infty \frac{b^p}{\Gamma(p)}\, x^{P-1} e^{-bx}\ dx = \frac{b^p}{\Gamma(p)} \int_0^\infty x^{P-1} e^{-bx}$$

and $f(x)$ is non-negative function.

The characteristic function of gamma distribution:

We have $\emptyset(t) = \int_{-\infty}^\infty e^{itx} f(x) dx$

$$= \int_0^\infty\ e^{itx} \frac{b^p}{\Gamma(p)}\ x^{P-1} e^{-bx} dx$$

$$i.e., \emptyset(t) = \frac{b^p}{\Gamma(p)} \int_0^\infty x^{p-1} e^{-(b-it)x}\, dx \quad \ldots (1)$$

We know that, $\frac{\Gamma(p)}{a^p} = \int_0^\infty y^{p-1}\, e^{-ay} dy$ is valid when $a = b + ic$ where $b > 0$,

$\int_0^\infty x^{p-1}\, e^{-(b-it)x} dx = \frac{\Gamma(p)}{(b-it)^p}$

$$(1) \Rightarrow \emptyset(t) = \frac{b^p}{\Gamma(p)} \times \frac{\Gamma(p)}{(b-it)^p} = \frac{b^p}{(b-it)^p} = \frac{1}{\left(1-\frac{it}{b}\right)^p}$$

$$i.e., \emptyset(t) = \frac{1}{\left(1-\frac{it}{b}\right)^p}$$

Now,

$$\emptyset(t) = \frac{1}{\left(1-\frac{it}{b}\right)^p} = \left(1 - \frac{it}{b}\right)^{-p}$$

$$\emptyset'(t) = -p\left(1 - \frac{it}{b}\right)^{-p-1}\left(\frac{-i}{t}\right)$$

$$\emptyset'(t) = \frac{p_i}{b\left(1-\frac{it}{b}\right)^{p+1}}$$

$$\emptyset''(t) = (-p)\big(-(p+1)\big)\left(1 - \frac{it}{b}\right)^{-p-2}\left(-\frac{i}{b}\right)^2$$

$$\emptyset''(t) = \frac{(i)^2 p(p+1)}{b^2\left(1-\frac{it}{b}\right)^{p+2}}$$

$$\emptyset^k(t) = \frac{p(p+1)\ldots\ldots\big(p+(k-1)\big)i^k}{b^k\left(1-\frac{it}{b}\right)^{p+k}}$$

$$\therefore\ \emptyset^k(0) = \frac{p(p+1)\ldots\ldots\big(p+(k-1)\big)i^k}{b^k}$$

$$m_k = \frac{\emptyset^k(0)}{i^k} = \frac{p(p+1)\ldots\ldots\big(p+(k-1)\big)i^k}{b^k i^k}$$

$$\therefore m_k = \frac{p(p+1)\ldots\ldots\big(p+(k-1)\big)}{b^k}.$$

In particular, we have

$$m_1 = \frac{p}{b}, m_2 = \frac{p(p+1)}{b^2}, \mu_2 = \frac{p}{b^2}$$

Example.

The random variable X has the gamma distribution with the density given by the formula

$$F(x) = \begin{cases} 0 \ for\ x \leq 0 \\ 2e^{-2x} \quad for\ x > 0 \end{cases}$$

What is the probability that X is not smaller than two?

Solution.

$$P(X \geq 2) = \int_2^\infty f(x)dx$$

$$= \int_2^\infty 2e^{-2x}\,dx = e\left(\frac{e^{-2x}}{-2}\right)_2^\infty$$

$$= (-0 + e^{-4}) \cong 0.0183$$

Definition.

The random variable with density $f(x)$, defined by

$$f(x) = \begin{cases} 0 & for\ x \leq 0 \\ \lambda e^{-\lambda x} & for\ \ x > 0 \end{cases}$$

where, $\lambda > 0$, has an ***exponential distribution***.

Theorem 4.8. (Addition theorem for random variable with gamma distribution)

Let $X = X_1 + X_2$, where X_1, X_2 are independent variables with gamma distribution then X is also has the gamma distribution.

Proof.

Let X_1 & X_2 be two independent random variables with gamma distribution.

Let $\emptyset_1(t), \emptyset_2(t)$ are the characteristic function X_1 and X_2 respectively.

i.e., $\emptyset_!(t) = \frac{1}{\left(1-\frac{it}{b}\right)^{p_1}}; \emptyset_2(t) = \frac{1}{\left(1-\frac{it}{b}\right)^{p_2}}$

Let $\emptyset(t)$ be the characteristic function of X

$\emptyset(t) = \emptyset_1(t)\emptyset_2(t)$

$$= \frac{1}{\left(1-\frac{it}{b}\right)} \cdot \frac{1}{\left(1-\frac{it}{b}\right)} \qquad (\because X_1 \,\&\, X_2 \; are \; independent)$$

$$\emptyset(t) = \frac{1}{\left(i-\frac{it}{b}\right)^{p_1+p_2}}$$

$\therefore X$ has the gamma distribution.

Theorem 4.9.

Let the independent random variables X and Y with non-independent distributions take on only positive values. Then X and Y have the gamma distribution with the same parameter b iff the random variables U and V, where $U = X + Y; V = \frac{X}{Y}$ are independent.

4.9. The Beta distribution

sNote that,

1. $\beta(p, q) = \int_0^1 x^{p-1}(1-x)^{q-1}\, dx \quad where\; p > 0, q < 0$
2. $\beta(p, q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$

Definition.

The random variable X has a ***beta distribution*** if its density is

$$f(x) = \begin{cases} \frac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1} \; for \; 0 < x < 1 \\ 0 \qquad\qquad for \; x \leq 0 \; and \; x \geq 1 \end{cases}$$

where, $p > 0, q < 0$.

Theorem 4.10.

Find the moments of the beta distribution

Proof.

$$m_k = \int x^k f(x) dx$$

$$m_k = \int_0^1 x^k \frac{1}{\beta(p,q)} x^{p-1}(1-x)^{q-1} \, dx$$

$$= \frac{1}{\beta(p,q)} \int_0^1 x^{p+k-1}(1-x)^{q-1} \, dx$$

$$= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \times \beta(p+k, q)$$

$$= \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \times \frac{\Gamma(p+k)\,\Gamma(q)}{\Gamma(p+k+q)}$$

$$= \frac{\Gamma(p+q)\,\Gamma(p+k)}{\Gamma(p)\,\Gamma(p+q+k)}$$

$$= \frac{\Gamma(p+q)\, p\,(p+1)(p+2)\ldots\ldots\ldots(p+k-1)\Gamma(p)}{\Gamma(p)(p+q+k-1)\ldots\ldots\ldots(p+q+2)(p+q+1)(p+q)\Gamma(p+q)}$$

$$\therefore m_k = \frac{P(p+1)\ldots\ldots..(p+k-1)}{(p+q)(p+q+1)\ldots\ldots..(p+q+k-1)}$$

In particular,

$$m_1 = \frac{p}{p+1}; m_2 = \frac{p(p+1)}{(p+q)(p+q+1)}$$

$$\mu_2 = m_2 - m_1^2$$

$$= \frac{p(p+1)}{(p+q)(p+q+1)} - \frac{p^2}{(p+q)^2}$$

$$= \frac{p(p+1)(p+q)-p^2(p+q+1)}{(p+q)^2(p+q+1)}$$

$$= \frac{(p^2+p)(p+q)-p^3-p^2q-p^2}{(p+q)^2(p+q+1)}$$

$$= \frac{p^3+p^2q+p^2+pq-p^3-p^2q-p^2}{(p+q)^2(p+q+1)}$$

Remark.

The density of the beta distribution with $p = q = 2$ represent as follow as in note 1 given above.

Example.

The random variable X has the beta distribution with $p = q = 2$; hence its density $f(x)$ is

$$f(x) = \begin{cases} 0 & for\ x \leq 0\ \ x \geq 1 \\ \frac{\Gamma(4)}{\Gamma(2)\Gamma(2)} x(1-x) & for\ 0 < x < 1 \end{cases}$$

i.e., $f(x) = \begin{cases} 0 & for\ x \leq 0\ x \geq 1 \\ 6x(1-x) & for\ 0 < x < 1 \end{cases}$ .What is the probability that X is not greater than 0.2?

Solution.

$$p(x \leq 0.2) = \int_0^{0.2} f(x)dx$$

$$= \int_0^{0.2} 6x(1-x)dx$$

$$= 6\ \left[\frac{x^2}{2} - \frac{x^3}{3}\right]_0^{0.2}$$

$$= 6\left[\frac{(0.2)^2}{2} - \frac{(0.2)^3}{3}\right]$$

$$P(X \leq 0.2) = 0.104.$$

4.10. The Cauchy and Laplace distributions

Definition.

The random variable X has a ***Cauchy distribution*** if its density is

$$f(x) = \frac{1}{\pi}\frac{\lambda}{\lambda^2+(x-\mu)^2} \text{ ,where } \lambda > 0$$

The function $f(x)$ is non-negative.

By substituting $y = \frac{x-\mu}{\lambda}$, we obtain

$$\int_{-\infty}^{\infty} f(x)dx = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{\lambda^2+(x-\mu)^2}$$

$$= \frac{1}{\lambda^2}\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dx}{1+\left(\frac{x-\mu}{\lambda}\right)^2}$$

$$= \frac{1}{\lambda^2}\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{\lambda dy}{(1+y^2)} \qquad \left(\text{since } y = \frac{x-\mu}{\lambda}\right.$$

$$dy = \frac{dx}{\lambda}$$

$$\left.\Rightarrow dx = \lambda dy\right)$$

$$= \frac{1}{\pi}\left[\frac{1}{\lambda}\int_{-\infty}^{\infty}\frac{dy}{(1+y^2)}\right]$$

$$= \frac{1}{\pi}[\tan^{-1} y]_{\infty}^{\infty}$$

$$= \frac{1}{\pi}\left[\frac{\pi}{2} + \frac{\pi}{2}\right]$$

$$= \frac{1}{\pi} \times \pi = 1$$

$$\therefore \int_{-\infty}^{\infty} f(x)dx = 1$$

The characteristic function of *Y*

Y has the density function

$$f(y) = \frac{1}{\pi} \frac{1}{1+y^2}$$

The characteristic function of Y is

$$\emptyset(t) = \int_{-\infty}^{\infty} e^{ity} f(y)dy$$

$$\emptyset(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{ity} \frac{1}{1+y^2} dy \dots . (I)$$

Consider the first density function $f_1(y) = \frac{1}{2} e^{-|y|} \rightarrow (1)$

Find the characteristic function for this density function is

$$\emptyset_1(t) = \int_{-\infty}^{\infty} e^{ity} f_1(y)dy$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} e^{ity} e^{-|y|} dy$$

$$= \frac{1}{2} \int_{-\infty}^{\infty} (costy + isinty) e^{-|y|} dy$$

$$\emptyset_1(t) = \frac{1}{2} \times 2 \int_0^{\infty} costy\, e^{-y} dy \quad (\because even\ function)$$

$$= [-e^{-y}]_0^{\infty} - t \int_0^{\infty} sinty\, e^{-y}\, dy$$

$$= 1 - t \int_0^{\infty} sinty\, e^{-y} dy$$

$$= 1 - t \left\{ [-e^{-y} sinty]_0^{\infty} + t \int_0^{\infty} e^{-y} costydy \right\}$$

$$= 1 - t^2 \int_0^{\infty} e^{-y} costydy$$

$$\emptyset_1(t) = \int_0^{\infty} e^{-y} costydy = 1 - t^2 \int_0^{\infty} e^{-y} costydy$$

$$\Rightarrow t^2 \int_0^{\infty} e^{-y} costydy + \int_0^{\infty} e^{-y} costydy = 1$$

$$\Rightarrow (1 + t^2) \int_0^{\infty} e^{-y} costydy = 1$$

$$\Rightarrow \int_0^{\infty} e^{-y} costydy = \frac{1}{1+t^2}$$

$$\Rightarrow \emptyset_1(t) = \frac{1}{1+t^2}$$

The density is

$$f_1(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ity}\, \emptyset_1(t) dt$$

$$f_1(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ity}}{1+t^2} dt \quad \rightarrow (2)$$

From (1) & (2),

$$\frac{1}{2} e^{-|y|} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-ity}}{1+t^2} dt$$

$$e^{-|y|} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-ity}}{1+t^2} dt$$

Changing e^{-ity} into e^{ity} under the integral sign (this doesnot affect the value of the integral) and changing the roles of t and y, we obtain

$$e^{-|t|} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{e^{-ity}}{1+y^2} dy \quad \rightarrow (3)$$

The R.H.S of (I) and (3) are same.

$$\therefore \emptyset(t) = e^{-|t|}$$

Since X is a linear transformation of Y, for the characteristic function $\emptyset_2(t)$ of X we obtain the formula

$$\emptyset_2(t) = e^{i\mu t - \lambda|t|}$$

Theorem 4.11. (Addition theorem for the Cauchy distribution).

Let X_1 and X_2 be two independent random variables with Cauchy distribution then $X = X_1 + X_2$ also has Cauchy distribution.

Proof.

Let X_1 and X_2 be two independent random variables with densities.

$$g_1(x) = \frac{1}{\pi}\frac{\lambda_1}{\lambda_1^2+(x-\mu_1)^2}; \quad g_2(x) = \frac{1}{\pi}\frac{\lambda_2}{\lambda_2^2+(x-\mu_2)^2}\ (\lambda_1, \lambda_2 > 0)$$

The characteristic function of X_1 and X_2 are

$$\psi_1(t) = e^{i\mu_1 t-\lambda_1|t|}; \psi_2(t) = e^{i\mu_2 t-\lambda_2|t|}$$

respectively.

Consider the random variable $X = X_1 + X_2$

Let $\psi(t)$ be the characteristic function of X. Then

$$\psi(t) = \psi_1(t)\,\psi_2(t) \quad (since, X_1\ \&\ X_2\ are\ independent)$$

$$= e^{i\mu_1 t-\lambda_1|t|} . e^{i\mu_2 t-\lambda_2|t|}$$

$$\psi_1(t) = e^{i(\mu_1+\mu_2)t-(\lambda_1+\lambda_2)\,|t|}$$

which is also characteristic function of ***Cauchy distribution***.

Remark.

X has a ***Laplace distribution*** if $X = \lambda Y + \mu,$ where Y has the density function $f_1(y) = \frac{1}{2}e^{-|y|}$.

∴ The density function of X is

$$f(x) = \frac{1}{2\lambda}\, e^{\left(\frac{-|x-\mu|}{\lambda}\right)}\ (\lambda > 0)$$

The characteristic function of X is

$$\emptyset(t) = \frac{e^{i\mu t}}{1+\lambda^2 t^2}$$

Random variable with a Laplace distribution has moments of any order.

UNIT – V

LIMITS THEOREMS

5.1. Stochastic Convergence

Example 1

The random variable Y_n can take on the value $0, \frac{1}{n}, \frac{2}{n}, \ldots.., \frac{n-1}{n}, 1$ and its probability function is given by the formula

$P\left(Y_n = \frac{r}{n}\right) = \binom{n}{r}\frac{1}{2^n}$ $(r = 0,1, \ldots.., n)$

Consider the random variable X_n defined by the formula

$X_n = Y_n - \frac{1}{2}$.

Thus X_n can take on the values

$-\frac{1}{2}, \frac{2-n}{2n}, \frac{4-n}{2n}, \ldots.., \frac{n-4}{2n}, \frac{n-2}{2n}, \frac{1}{2}$

The probability function of X_n is given by the formula

$$P\left(X_n = \frac{2r-n}{2n}\right) = \binom{n}{r}\frac{1}{2^n}$$

Let $n = 2$. The random variable can take on the values

$-0.5, 0, 0.5$

with respective probabilities

$\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$.

Let ε be a positive number, say $\varepsilon = 0.3$

$P(|X_2| > 0.3 = P\left(X_2 = \frac{-1}{2}\right) + P\left(X_2 = \frac{1}{2}\right) = 0.5$

Let $n = 5$. The random variable X_5 can take on the values

$-0.5, -0.3, -0.1, 0.1, 0.3, 0.5$

with respective probabilities

$$\frac{1}{32}, \frac{5}{32}, \frac{10}{32}, \frac{10}{32}, \frac{5}{32}, \frac{1}{32}$$

Hence

$$P(|X_5|) > 0.3 = 0.0625$$

Let $n = 10.$ The random variable X_{10} can take on the values

$$-0.5, -0.4, -0.3, -0.2, -0.1, 0.0, 0.1, 0.2, 0.3, 0.4, 0.5$$

with respective probabilities

$\frac{1}{1024}, \frac{10}{1024}, \frac{45}{1024}, \frac{120}{1024}, \frac{210}{1024}, \frac{252}{1024}, \frac{120}{1024}, \frac{45}{1024}, \frac{10}{1024}, \frac{1}{1024}$

$P(|X_{10}| > 0.3 \cong 0.02$.

For $n = 10,$ the probability that X_n will exceed $\varepsilon = 0.3/n$ absolute value is very small.

Definition.

The sequence $\{X_n\}$ of random variable is called ***stochastically convergent to zero*** if for every $\varepsilon > 0$ the relation

$$\lim_{n\to\infty} P(|X_n| > \varepsilon = 0$$

is satisfied.

Theorem 5.1.

Let $F_n(x)(n = 1,2,3, \dots.)$ be the distribution function of the random variable X_n. The sequence $\{X_n\}$ is stochastically convergent to zero iff the sequence $\{F_n(x)\}$ satisfies the relation

$$\lim_{n\to\infty} F_n(x) = \begin{cases} 0 & for\ x \le 0 \\ 1 & for\ x > 0 \end{cases}$$

Proof.

Suppose that the sequence $\{X_n\}$ is stochastically convergent to zero.

$\therefore \lim_{n\to\infty} P(|X_n| > \varepsilon = 0 \quad \to (1)\ \ \forall\varepsilon > 0$

For every $\varepsilon > 0$,

$\lim_{n\to\infty} P(X_n < \varepsilon) = 0$

$\lim_{n\to\infty} F_n(-\varepsilon) = 0 \quad \to (A)$

$\lim_{n\to\infty} P(X_n > \varepsilon) \ = 1 - P(X_n < \varepsilon) - P(X_n = \varepsilon)$

$P(X_n > \varepsilon) = 1 - P(X_n < \varepsilon) - P(X_n = \varepsilon)$

$P(X_n > \varepsilon) = 1 - F_n(\varepsilon) - P(X_n = \varepsilon) \ \to (2)$

Since, for every $\varepsilon > 0$, we can find an $\varepsilon_1$such that $0 < \varepsilon_1 < \varepsilon$

From (1), for an arbitrary $\varepsilon > 0$, we have

$\lim_{n\to\infty} P(X_n = \varepsilon) \ = 0 \quad \to (3)$

Substitute (3) in (2) we get

$$\lim_{n\to\infty} P(X_n > \varepsilon = \lim_{n\to\infty}\big(1 - F_n(\varepsilon) - P(X_n = \varepsilon)\big) = 0 \ = \lim_{n\to\infty}\big(1 - F_n(\varepsilon)\big)$$

$\Rightarrow \lim_{n\to\infty} F_n(\varepsilon) = 0 \qquad \to (B)$

Replace ε by $-x$ in (A) & $\varepsilon\ by\ x\ in\ (B)$ where $x > 0$, We get

$\lim_{n\to\infty} F_n(+x) = 0$ and $\lim_{n\to\infty} F_n(x) = 1$

$\therefore \lim_{n\to\infty} F_n(x) = \begin{cases} 0 & for\ x \le 0 \\ 1 & for\ x > 0 \end{cases}$

Conversely,

Suppose, $\lim_{n\to\infty} F_n(x) = \begin{cases} 0 & for\ x \le 0 \\ 1 & for\ x > 0 \end{cases}$

Then for arbitrary $\varepsilon > 0$ we have

$\lim_{n\to\infty} P(X_n < -\varepsilon) = \lim_{n\to\infty} F_n(-\varepsilon) \ = 0$

$\therefore \lim_{n\to\infty} P(X_n < -\varepsilon) \ = 0$

$\Rightarrow \lim_{n\to\infty} P(-X_n > \varepsilon) = 0$

$\lim_{n\to\infty} P(X_n > \varepsilon) \le \lim_{n\to\infty} P(X_n \ge \varepsilon)$

$= \lim_{n\to\infty} [1 - F_n(\varepsilon)]$

$= 1 - \lim_{n\to\infty} F_n(\varepsilon)$

$= 1 - 1$

$= 0$

$\therefore \lim_{n\to\infty} P(X_n > \varepsilon) = 0$

$\Rightarrow \lim_{n\to\infty} P(|X_n| \ > \varepsilon) = 0$

Hence the proof.

Remark.

The random variable X with a one-point distribution such that $P(X = 0) = 1$ has the distribution function $F(x) = \begin{cases} 0 & for\ x \leq 0 \\ 1 & for\ x > 0 \end{cases}$

This distribution function is continuous at every point $x \neq 0$.

From above theorem.

For arbitrary $\varepsilon > 0$

Since $\lim_{n\to\infty} P(X_n < -\varepsilon) = \lim_{n\to\infty} F_n(-\varepsilon)$

$= 0$

$\therefore \lim_{n\to\infty} P(X_n < -\varepsilon) = 0$

$\lim_{n\to\infty} P(X_n > \varepsilon) = 1 - \lim_{n\to\infty} F_n(\varepsilon)$

$= 1 - 1$

$\lim_{n\to\infty} P(X_n > \varepsilon) = 0$

$\therefore \lim_{n\to\infty} P(|X_n| > \varepsilon) = 0$

$\therefore$ For every point $x \neq 0$ the sequence of distribution function $F_n(x)$ converges to the distribution function $F(x)$.

i.e, The sequence of distribution function $F_n(x)$ of random variable convergent stochastically to zero, converges to the distribution function of the one-point distribution at every point $x \neq 0$.

Since the points $x \neq 0$ are continuity points of this distribution function, we can formulate the preceding result in the following way:

The sequence $\{X_n\}$ of random variable is stochastically convergent to zero iff the sequence $\{F_n(x)\}$ of distribution functions of these random variable is convergent *to the distribution function* $F(x)$ *given by* $F(x) = \begin{cases} 0 & for\ x \leq 0 \\ 1 & for\ x > 0 \end{cases}$ *at every continuity point of the latter.*

Note:

1. The fact that at the point of discontinuity of $F(x)$. That is, at the point $x = 0$, the sequence $\{F_n(0)\}$ may not converge to $F(0)$.
2. The sequence of random variables$\{Y_n\} = \{X_n - c\}$ is stochastically convergent to zero.
3. The sequence of random variables$\{Z_n\} = \{X_n - X\}$ is stochastically convergent to zero.

5.2. Bernoulli's Law of Large numbers

Theorem 5.2 (Bernoulli law of large numbers)

Let $\{Y_n\}$ be the sequence of random variable with probability functions given by

$$P\left(Y_n = \frac{r}{n}\right) = \binom{n}{r} p^r (1-p)^{n-r} \rightarrow (1)$$

where, $0 < p < 1$ and r can take on the values $0,1,2, \dots \dots n$

Let $X_n = Y_n - p \quad \rightarrow (2)$

The sequence of random variable $\{X_n\}$ given by (1) and (2) is stochastically convergent to 0.

i.e, for every $\varepsilon > 0$, $\lim_{n\to\infty} P(|X_n| > \varepsilon) = 0$.

Proof.

The Chebyshev inequality is

$P(|X - m_1| \geq k\sigma \leq \frac{1}{K^2}$

Put $E(X_n) = 0$ & $\sigma_n = \sqrt{D^2(X_n)}$

$$= \sqrt{\frac{P(1-P)}{n}}$$

$P(|X_n - m_1| \geq k\sigma) \leq \frac{1}{k^2}$

i.e, $P\left(|X_n| > k\sqrt{\frac{P(1-P)}{n}}\right) \leq \frac{1}{k^2}$

where k is an arbitrary positive number.

Let $k = \varepsilon\sqrt{\frac{n}{p(p-1)}}$

$P\left(|X_n| > \varepsilon\sqrt{\frac{n}{p(p-1)}}\sqrt{\frac{P(1-P)}{n}}\right) \leq \frac{p(p-1)}{n\varepsilon^2}$

$P(|X_n| > \varepsilon) \leq \frac{p(p-1)}{n\varepsilon^2} <= \frac{1}{n\varepsilon^2}$

$\Rightarrow \lim_{n\to\infty} P(|X_n| > \varepsilon) = 0$ for every $\varepsilon > 0$

5.3. The Convergence of a Sequence of Distribution Functions

Definition.

The sequence $\{F_n(x)\}$ of distribution function of the random variables $\{X_n\}$ is called ***convergent***, if there exist, a distribution function $F(x)$ such that, at every continuity point of $F(x)$, the relation

$$\lim_{n\to\infty} F_n(x) = F(x)$$

is satisfied. The distribution function $F(x)$ is called the ***limit distribution function***.

Remark.

Consider the example (1), By theorem 6.3.1, the sequence $\{X_n\}$ of random variable defined by $X_n = Y_n - \frac{1}{2}$ is stochastically convergent to zero.

∴ The sequence $\{F_n(x)\}$ of their distribution function converges to the distribution function $F(x)$ defined by $F(x) = \begin{cases} 0 & for\ x \leq 0 \\ 1 & for\ x > 0 \end{cases}$

This distribution function is discontinuous at $x = 0$.

i.e., $\{F_n(0)\}$ is not convergent to $F(0)$.

consider the subsequence of the sequence $\{F_n(0)\}$ containing only terms with the odd indices $n = 2k + 1$. The random variable X_{2n+1} can take on the values

$$-\frac{1}{2}, \frac{2-(2k+1)}{2(2k+1)}, \frac{4-(2k+1)}{2(2k+1)}, \ldots\ldots, \frac{2k+1-4}{2(2k+1)}, \frac{2k+1-2}{2(2k+1)}, \frac{1}{2}$$

For every k, half of these terms are each less than zero, the other half greater than zero. The probability that X_{2k+1} will take on a value less than zero equals 0.5

∴ For every k we have $P(X_{2k+1} < 0) = F_{2k+1}(0) = 0.5$

Since, $F(0) = 0$, we have

$\lim_{k\to\infty} F_{2k+1}(0) = 0.5 \neq F(0) \quad \to (1)$

From (1) it follows that, $\lim_{n\to\infty} F_n(0) \neq F(0) \qquad \to (1)$

From (1) it follows that, $\lim_{n\to\infty} F_n(0) \neq F(0)$.

Example 1.

Let us consider the sequence $\{X_n\}$ of random variable with the one-point distributions given by the formula

$$P(X_n = n) = 1 \ (n = 1,2, \dots \dots .)$$

The distribution function $F_n(x)$ of X_n is of the form

$$F_n(x) = \begin{cases} 0 & for\ x \le n \\ 1 & for\ x > n \end{cases}$$

We have the relation

$$\lim_{n\to\infty} F_n(x) = 0 \qquad\qquad (-\infty < x < \infty)$$

∴ The sequence $\{F_n(x)\}$ convergent to 0

i.e., The sequence $\{F_n(x)\}$ is not convergent to a distribution function $F_n(x)$.

Remark

1. Let the sequence $\{F_n(x)\}$ be convergent to the distribution function $F(x)$. Let a and b, where $a < b$, be two arbitrary continuity points of the limit distribution function $F(x)$. Then we have,
   $\underset{n\to\infty}{lim} P(a \le X_n < b) = F(b) - F(a)$

For,

$P(a \le X_n < b) = F_n(b) - F_n(a)$

$$\lim_{n\to\infty} P(a \le X_n < b) = \underset{n\to\infty}{lim} [F_n(b) - F_n(a)]$$

$$= \underset{n\to\infty}{lim} F_n(b) - \underset{n\to\infty}{lim} F_n(a)$$

$$= F(b) - F(a)$$

(Since, a and b are continuity points of distribution $F(x)$ i.e., $F_n(b) \to F(b)$ & $F_n(a) \to F(a)$)

2. Let the sequence $\{F_n(x)\}$ be convergent to the distribution function $F(x)$. Let $P_n(S)$ and $P(S)$ denote the probability function corresponding respectively to the distribution function $F_n(x)$ and $F(x)$. Then we have $\lim_{n\to\infty} P_n(S) = P(S)$.

Example 2.

The random variable $X_n(n = 1,2,3, \dots \dots)$ has the density $f_n(x)$ given by

$$f_n(x) = \begin{cases} \frac{2^n}{\varepsilon} & if \frac{i}{n} - \frac{\varepsilon}{ne^n} < x < \frac{i}{n} (i = 1,2, \dots \dots, n) \\ 0 & otherwise \end{cases}$$

where, $0 < \varepsilon < 1$. The distribution function $F_n(x)$ of X_n is

$$F_n(x) = \begin{cases} 0 \quad if\ x \le 0 \\ \frac{i-1}{n} \quad if\ \frac{i-1}{n} \le x \le \frac{i}{n} - \frac{\varepsilon}{n2^n} \\ \frac{i-1}{n} + \frac{2^n\left(x - \frac{i}{n} + \frac{\varepsilon}{n2^n}\right)}{\varepsilon} \quad if\ \frac{i}{n} - \frac{\varepsilon}{n2^n} < x < \frac{i}{n} \\ 1 \quad if\ x \ge 1 \end{cases}$$

Thus for every x in the interval $I = [0,1]$ we have

$$0 \le x - F_n(x) \le \frac{1}{n}$$

By considering the values taken by $F_n(x)$ outside the interval I, we obtain for every real x

$$\lim_{n\to\infty} F_n(x) = F(x) = \begin{cases} 0 & for\ x \le 0 \\ x & for\ 0 < x < 1 \\ 1 & if\ x \ge 1 \end{cases}$$

Definition.

The sequence of distribution function $\{F_n(x_1, \dots \dots, x_n)\}$ of random vector $\left(X_{n_1}, X_{n_2}, \dots \dots, X_{n_k}\right)$ is ***convergent*** if there exist a distribution function $F(x_1, x_2, \dots \dots, x_k)$ such that at every one of its continuity points.

$$\lim_{n\to\infty} F_n(x_1, x_2, \dots \dots, x_k) = F(x_1, x_2, \dots \dots, x_k)$$

Theorem 5.2.

Let $\{F_n(x_1, x_2, \dots \dots, x_k)\}$ $(n = 1,2, \dots \dots)$ be a sequence of distribution functions of random vectors $\left(X_{n_1}, X_{n_2}, \dots \dots, X_{n_k}\right)$ and let $F(x_1, x_2, \dots \dots, x_k)$ and $P(S)$ be the distribution function and probability function of a random vector $(X_1, X_2, \dots \dots, X_k)$ respectively. $\lim_{n\to\infty} F_n(x_1, x_2, \dots \dots, x_k) = F(x_1, x_2, \dots \dots, x_k)$ holds iff for every function $2(x_1, \dots \dots, x_k)$ continuous on a set S satisfying the relation $P(S) = 1$

$\lim_{n\to\infty} H_n(\alpha) = H(\alpha)$ holds at every continuity point α of $H(\alpha)$ where $H_n(\alpha)$ and $H(\alpha)$ are the distribution function of $g\left(X_{n_1}, X_{n_2}, \dots \dots, X_{n_k}\right)$ and $g(x_1, x_2, \dots \dots, x_k)$ respectively.

5.4. The De Moivre-Laplace Theorem

Let $\{X_n\}$ be a sequence of random variables with the binomial distribution. For ever n the random variable X_n can take on the values $0,1, \dots \dots, n$ and its probability is

$$P(X_n = r) = \binom{n}{r} p^r q^{n-r}$$

where $0 < p < 1$ and $q = 1 - p$

Clearly we have, $E(X_n) = np; D^2(X_n) = npq$

Consider the sequence $\{Y_n\}$ of standardizer random variables $Y_n = \frac{X_n - np}{\sqrt{npq}}$.

Theorem 5.3.(De Moivre-Laplace theorem)

Let $\{F_n(y)\}$ be the sequence of distribution functions of the random variables $Y_n = \frac{X_n - np}{\sqrt{npq}}$, where the X_n have the binomial distribution given by $P(X_n = r) = \binom{n}{r} p^r q^{n-r}$. If $0 < p < 1$, then for every y we have the relation $\lim_{n\to\infty} F_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{y^2}{2}} dy$.

Proof.

Given $P(X_n = r) = \binom{n}{r} p^r q^{n-r}$

The characteristic function $\emptyset_x(t)$ of X_n is

$$\emptyset_x(t) = \left(q + pe^{it}\right)^n$$

Given $Y_n = \frac{X_n - np}{\sqrt{npq}}$

The characteristic function $\emptyset_y(t)$ of the random variable Y_n is

$$\emptyset_y(t) = e^{\frac{-npit}{\sqrt{npq}}} \left(q + pe^{\frac{it}{\sqrt{npq}}}\right)^n$$

$$= \left(e^{\frac{-pit}{\sqrt{npq}}}\right)^n \left(q + pe^{\frac{it}{\sqrt{npq}}}\right)^n$$

$$= \left[qe^{\frac{-pit}{\sqrt{npq}}} + pe^{(1-p)\frac{it}{\sqrt{npq}}}\right]^n$$

$$\emptyset_y(t) = \left[qe^{\frac{-pit}{\sqrt{npq}}} + pe^{\frac{pit}{\sqrt{npq}}}\right]^n \quad \rightarrow (1)$$

Let us expand the function e^{iz} in the ngd of $z = 0$ according to the Taylor formula for k terms with the remainder in the peano form,

$$e^{iz} = \sum_{j=0}^{k} \frac{(iz)^j}{j!} + 0(z^k)$$

We obtain

$$pe^{\frac{pit}{\sqrt{npq}}} = p + it\sqrt{\frac{pq}{n}} - \frac{qt^2}{2n} + 0\left(\frac{t^2}{n}\right) \quad \rightarrow (2)$$

$$qe^{\frac{-pit}{\sqrt{npq}}} = q - it\sqrt{\frac{pq}{n}} - \frac{pt^2}{2n} + 0\left(\frac{t^2}{n}\right) \rightarrow (3) \quad (\because expalined\ in\ the\ class)$$

where for every t we have,

$$lim_{n\to\infty} n0\left(\frac{t^2}{n}\right) = 0 \quad \rightarrow (4)$$

Sub (2) +(3) in (1)

$$\emptyset_y(t) = \left[q - it\sqrt{\frac{pq}{n}} - \frac{pt^2}{2n} + 0\left(\frac{t^2}{n}\right) + o + it\sqrt{\frac{pq}{n}} - \frac{qt^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

$$= \left[p + q - \frac{(p+q)t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

$$\emptyset_y(t) = \left[1 - \frac{t^2}{n} + 0\left(\frac{t^2}{n}\right)\right]^n$$

$$\log \emptyset_y(t) = n \log\left[1 - \frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right] = nlog(1 + z)$$

For every fixed t for sufficiently large n, we have $|z| < 1$

$$\therefore \log \emptyset_y(t) = \frac{-t^2}{2n} + n0\left(\frac{t^2}{n}\right)$$

$$lim_{n\to\infty} \log \emptyset_y(t) = lim_{n\to\infty}\left(\frac{-t^2}{2} + m0\left(\frac{t^2}{n}\right)\right)$$

$$= \frac{-t^2}{2}\ lim_{n\to\infty} n0\left(\frac{t^2}{n}\right)$$

$$\therefore lim_{n\to\infty} \log \emptyset_y(t) = e^{\frac{-t^2}{2}} \quad \big(using\ (4)\big)$$

$$lim_{n\to\infty} \emptyset_y(t) = e^{-\frac{t^2}{2}}$$

$\therefore$ The sequence of characteristic function $\emptyset_y(t)$ of the standardized random variables $Y_n = \frac{X_n - np}{\sqrt{npq}}$ convers as $n \to \infty$ to the characteristic function of a random variable with a normal distribution whose distribution function is $\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{y^2}{2}}\ dy$

$$\therefore lim_{n\to\infty} F_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{y^2}{2}}\ dy$$

Hence proved.

Remark.

Let y_1 and y_2 be two arbitrary points with $y_1 < y_2$.

We know that, $lim_{n\to\infty} F_n(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{y^2}{2}}\ dy \to (1)$

From the above relation,

$$lim_{n\to\infty} P(y_1 < Y < y_2) = lim_{n\to\infty}[F_n(y_2) - F_n(y_1)]$$

$$= lim_{n\to\infty} F_n(y_2) - lim_{n\to\infty} F_n(y_1)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_2} e^{-\frac{y^2}{2}}\ dy - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y_1} e^{-\frac{y^2}{2}}\ dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y_2} e^{-\frac{y^2}{2}}\ dy$$

$$\therefore lim_{n\to\infty} P(y_1 < Y < y_2) = \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y_2} e^{-\frac{y^2}{2}}\ dy \ \to (2)$$

$P(y_1 < Y < y_2) = P\left(y_1 < \frac{X_n - np}{\sqrt{npq}} < y_2\right)$

$$= P\left(y_1\sqrt{npq} + np < X_n < y_2\sqrt{npq} + np\right)$$

$\therefore lim_{n\to\infty} P\left(y_1\sqrt{npq} + np < X_n < y_2\sqrt{npq} + np\right) = \frac{1}{\sqrt{2\pi}} \int_{y_1}^{y_2} e^{-\frac{y^2}{2}}\, dy$

Let $x_1 = y_1\sqrt{npq} + np; x_2 = y_2\sqrt{npq} + np \quad \to (3)$

$P(x_1 < X_n < x_2) \cong \int_{y_1}^{y_2} e^{-\frac{y^2}{2}}\, dy$ where y_1 and y_2 are determined by (3)

We say that the random variable X_n has a ***asymptotically normal distribution*** $N\left(np; \sqrt{npq}\right)$.

Replacing y_1 and y_2 with

$y_1 + \frac{1}{2\sqrt{npq}}$ and $y_2 - \frac{1}{2\sqrt{npq}}$ respectively we get a better approximation.

Example 1

We throw a coin $n = 100$times. We assign the number 1 to the appearance of heads and the number 0 to the appearance of tails. The probability of each of these events is equal to $p = q = 0.5$. what is the probability that heads will appear more than 50 times and less than 60 times?

Solution.

The random variable X_n take on values from $0\ to\ 100$.

Given, $n = 100; p = 0.5; q = 0.5$

$$E(X_n) = np = 100 \times 0.5 = 50; D^2(X_n) = npq$$

$$= 100 \times 0.5 \times 0.5$$

$$= 25$$

$\therefore E(X_n) = 50 \ \& \ D^2(X_n) = 25.$

We know that,

$$P(x_1 < X_n < x_2) \cong \frac{1}{\sqrt{2\pi}} \int_{y_!}^{y_2} e^{-\frac{t^2}{2}} \ dt$$

Where $x_1 = y_1\sqrt{npq} + np$ & $x_2 = y_2\sqrt{npq} + np$

Here, $x_1 = 50$ and $x_2 = 60$

$y_1 = \frac{50-50}{5}$ $\qquad$ $y_2 = \frac{60-50}{5}$

$y_1 = 0$ $\qquad$ $y_2 = 2$

$$P(50 < X_n < 60) \cong \frac{1}{\sqrt{2\pi}} \int_0^2 e^{-\frac{t^2}{2}} \ dt$$

$$\cong \frac{1}{\sqrt{2\pi}} \int_{0.1}^{1.9} e^{-\frac{t^2}{2}} \ dt$$

$$\cong \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1.9} e^{-\frac{t^2}{2}} \ dt - \int_{-\infty}^{0.1} e^{-\frac{t^2}{2}} \ dt$$

$$\cong \emptyset(1.9) - \emptyset(0.1)$$

$$= 0.97283 - 0.539828$$

$$= 0.431455$$

$P(50 < X_n < 60) \cong 0.4315.$

Remark.

1. From de Moivre-Laplace theorem we obtain theorem for the sequence of random variables

$$U_n = \frac{X_n}{n}$$

Where X_n has the binomial distribution given by

$$P(X_n = r) = \binom{n}{r} p^r q^{n-r}$$

Since, $E(U_n) = p$ & $D^2(U_n) = \frac{pq}{n}$

$$Z_n = \frac{U_n - P}{\sqrt{\frac{pq}{n}}} = \frac{\frac{X_n}{n} - p}{\sqrt{\frac{pq}{n}}} = \frac{X_n - np}{n\sqrt{\frac{pq}{n}}}$$

$$= \frac{X_n - np}{\sqrt{npq}} = Y_n$$

$\therefore Z_n = Y_n$

Since, the sequence $\{F_n(y)\}$ of distribution function of Y_n satisfies

$$lim_{n\to\infty} F_n = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{y^2}{2}}\ dy$$

$\therefore$ for the sequence $\{F_n(z)\}$ of the distribution function of z_n

$$lim_{n\to\infty} F_n(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}}\ dz$$

2. For every pair of constants z_1 and z_2 where $z_1 < z_2$.

$$lim_{n\to\infty} P\left(z_1 < \sqrt{\frac{n}{pq}}\ (U_n - p) < z_2\right) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}}\ dz$$

Let

$$u_1 = z_1\sqrt{\frac{pq}{n}} + p \qquad u_2 = z_2\sqrt{\frac{pq}{n}} + p \qquad \to (1)$$

$$P(u_1 < U_n < u_2) \cong \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}}\ dz \ \to (2)$$

Where z_1 and z_2 are determined by (1)

The random variable $U_n$satisfying the relation (2) has an asymptotically normal distribution $N\left(p;\ \sqrt{\frac{pq}{n}}\right)$.

Example

A box contains a collection of IBM cards corresponding to the workers from some branch of industry of the workers 20% are minors and 80% adults. We select one IBM card in a random way and mark the age given in this card. Before choosing the next card, we return the first one to the box, so that the probability of selecting the card corresponding to a minor remains 0.2. we observe ‘n’ cards in this manner. What value should ‘n’ have in order that the probability will be 0.95 that the frequency of cards corresponding to minor lies between 0.18 and 0.22?

Solution.

Let U_n be the frequency of the appearance of the card corresponding to a minor.

$E(U_n) = p$ and $D^2(U_n) = \frac{pq}{n}$

Here, $p = 0.2 \quad q = 1 - p = 1 - 0.2 = 0.8$

$E(U_n) = 0.2 \quad D^2(U_n) = \frac{0.16}{n}$

$\sqrt{D^2(U_n)} = \frac{0.4}{\sqrt{n}}$

$P(u_1 < U_n < u_2) \cong \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} \, dz$

Where $u_1 = z_1\sqrt{\frac{pq}{n}} + p \quad u_2 = z_2\sqrt{\frac{pq}{n}} + p$

Here, $u_1 = 0.18 \quad and \; u_2 = 0.22$

$0.18 = z_1\left(\frac{0.4}{\sqrt{n}}\right) + 0.2 \qquad and \;\; 0.22 = z_2\left(\frac{0.4}{\sqrt{n}}\right) + 0.2$

$$0.18 - 0.2 = z_1\left(\frac{0.4}{\sqrt{n}}\right) \qquad\qquad z_2 = (0022 - 0.2)\left(\frac{\sqrt{n}}{0.4}\right)$$

$$\frac{0.16}{0.4} \times \sqrt{n} = z_1 \qquad\qquad z_2 = \left(\frac{0.20}{0.4}\right)\sqrt{n}$$

$$z_1 = \frac{1.6}{4}\sqrt{n} \qquad\qquad z_2 = \frac{2}{4}\sqrt{n}$$

$z_1 = 0.4\sqrt{n}$ $\qquad\qquad\qquad\qquad$ $z_2 = 0.5\sqrt{n}$

$$P(0.18 < U_n < 0.22) \cong \frac{1}{\sqrt{2\pi}} \int_{0.4\sqrt{n}}^{0.5\sqrt{n}} e^{-\frac{z^2}{2}} \, dz \qquad \rightarrow (1)$$

We know that,

$$lim_{n\to\infty} P\left(z_1 < \sqrt{\frac{n}{pq}}(U_n - p) < z_2\right) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} \, dz$$

$$P(0.18 < U_n < 0.22) = p\left(\frac{-0.02}{\frac{0.4}{\sqrt{n}}} < \frac{U_n - 0.2}{\frac{0.4}{\sqrt{n}}} < \frac{0.02}{\frac{0.4}{\sqrt{n}}}\right)$$

$$= p\left(-0.05\sqrt{n} < \frac{U_n - 0.2}{0.4}\sqrt{n} < 0.05\sqrt{n}\right)$$

$$\cong 0.95$$

$$\therefore P(0.18 < U_n < 0.22) \cong 0.95 \qquad \rightarrow (2)$$

From (1) and (2)

$$\frac{1}{\sqrt{2\pi}} \int_{0.4\sqrt{n}}^{0.5\sqrt{n}} e^{-\frac{z^2}{2}} \, dz \cong 0.95$$

From normal table,

$0.5\sqrt{n} \cong 1.96$

$\sqrt{n} \cong \frac{1.96}{0.5}$

$n \cong 1537.$

5.5. The Lindeberg-Levy Theorem

Consider a sequence $\{X_k\}$ $(k = 1,2, \ldots\ldots)$ of equally distributed, independent random variables whose moment of the second order exists.

For every k denote

$E(X_k) = m; D^2(X_k) = \sigma^2$

Consider the random variable $\frac{1}{n}$ defined by

$$Y_n = X_1 + X_2 + \cdots \ldots + X_n$$

$E(Y_n) = nm$ and $D^2(Y_n) = n\sigma^2$

Let $Z_n = \frac{Y_n - mn}{\sigma\sqrt{n}}$…..(A).

Theorem 5.4. (Lindeberg – Levy theorem)

If $X_1, X_2, \ldots ..$ are independent random variables with the same distribution, whose standard deviation $\sigma \neq 0$ exists, then the sequence $\{F_n(z)\}$ of distribution functions of the random variables Z_n, given by formulas

$$z_n = \frac{Y_n - mn}{\sigma\sqrt{n}} \text{ and } Y_n = X_1 + X_2 + \cdots \ldots + X_n,$$

satisfies, for ever z, the equality

$$lim_{n\to\infty} F_n(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}}\, dz$$

Proof.

let $z_n = \frac{Y_n - mn}{\sigma\sqrt{n}}$ can be written in the form

$z_n = \frac{1}{\sigma\sqrt{n}} \sum_{k=1}^{n} (X_k - m)$

All the random variable $X_k - m$ have the same distribution, hence the characteristic function $\emptyset_z(t)$ of Z_n is

$$\emptyset_z(t) = \left[\emptyset_x\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n \rightarrow (1)$$

Assume that the existence of the first and second moments we have,

$E(X_k - m) = 0 \quad and\ D^2(X_k - m) = \sigma^2$

Expand the function $\emptyset_z(t)$ in a neighbourhood of the point $t = 0$ according to the MacLaurin formula

$$\emptyset_x(t) = 1 - \frac{1}{2}\sigma^2 t^2 + 0(t^2)$$

$$\emptyset_x\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 - \frac{1}{2}\sigma^2\frac{t^2}{\sigma^2 n} + 0\left(\frac{t^2}{\sigma^2 n}\right)$$

$$\emptyset_x\left(\frac{t}{\sigma\sqrt{n}}\right) = 1 - \frac{1}{2}\frac{t^2}{n} + 0\left(\frac{t^2}{n}\right) \qquad \rightarrow (2)$$

Substitute (2) in (1)

$$\emptyset_z(t) = \left[1 - \frac{1}{2}\frac{t^2}{n} + 0\left(\frac{t^2}{n}\right)\right]^n \qquad \rightarrow (3)$$

where for every t we have,

$$lim_{n\rightarrow\infty} n0\left(\frac{t^2}{n}\right) = 0 \qquad \rightarrow (4)$$

Let $u = -\frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)$

$$(2) \Rightarrow \emptyset_z(t) = [1+u]^n$$

$$\log \emptyset_z(t) = n\log(1+u)$$

$$= n\left[-\frac{t^2}{2n} + 0\left(\frac{t^2}{n}\right)\right]$$

$$\log \emptyset_z(t) = -\frac{t^2}{2} + n0\left(\frac{t^2}{n}\right)$$

$$lim_{n\rightarrow\infty} \log \emptyset_z(t) = lim_{n\rightarrow\infty}\left[-\frac{t^2}{2} + n0\left(\frac{t^2}{n}\right)\right]$$

$$= -\frac{t^2}{2} + lim_{n\rightarrow\infty} n0\left(\frac{t^2}{2}\right)$$

$lim_{n\to\infty} \log \emptyset_z(t) = -\frac{t^2}{2} \quad \big(using\ (4)\big)$

$lim_{n\to\infty} \emptyset_z(t) = e^{-\frac{t^2}{2}}$

$e^{-\frac{t^2}{2}}$ is the characteristic function of a random variable with the normal distribution.

[by theorem, if the sequence of characteristic function $\{\emptyset_n(t)\}$ converges at every point $t(-\infty < t < +\infty)$ to a function $\emptyset(t)$ continuous in same interval $|t| < c$, then the sequence $\{F_n(x)\}$ of corresponding distribution function converges to the distribution function $F(x)$ which corresponds to the characteristic function $\emptyset(t)$

$\therefore lim_{n\to\infty} F_n(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}}\ dz$

Remark.

Let z_1 and z_2 be two arbitrary numbers with $z_1 < z_2$. By relation in the last theorem we obtain

$$lim_{n\to\infty} P(z_1 < Z_n < z_2) = lim_{n\to\infty}[F_n(z_2) - F_n(z_1)]$$

$$= \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}}\ dz \ ....(5)$$

From formula (A) we obtain

$$p(z_1 < Z_n < z_2) = P\left(z_1 < \frac{Y_n - mn}{\sigma\sqrt{n}} < z_2\right)$$

$$= P\big(z_1\sigma\sqrt{n} + nm < Y_n < z_2\sigma\sqrt{n} + nm\big)$$

Thus, we obtain from formula (5)

$$lim_{n\to\infty} P\big(z_1\sigma\sqrt{n} + nm < Y_n < z_2\sigma\sqrt{n} + nm\big) = \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}}\ dz ...(6)$$

Let

$$y_1 = z_1\sigma\sqrt{n} + nm, \;\; y_2 = z_2\sigma\sqrt{n} + nm \ldots .(7)$$

Now we can write formula (6) in the asymptotic form

$$P(y_1 < Y_n < y_2) \cong \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} \, dz$$

where z_1 and z_2 are determined by relation (7). Thus, the random variable Y_n defined by formula $Y_n = X_1 + X_2 + \cdots \ldots + X_n$ has an asymptotically normal distribution $N(nm; \sigma\sqrt{n})$.

Example 1.

Suppose that the random variables $\{X_k\}(k = 1,2, \ldots \ldots)$ are independent and each of them has the same two – point distribution, i.e, for every k we have

$P(X_k = 1) = p; P(X_k = 0) = 1 - p$ where $0 < p < 1$

Consider the random variable $Y_n = X_1 + X_2 + \cdots \ldots + X_n$

$E(X_k) = p$ and $D^2(X_k) = pq$

By de Moivre-Laplace limit theorem that Y_n has an asymptotically normal distribution $N(np; \sqrt{npq})$.

Remark.

De Moivre-Laplace limit theorem is a particular case of Lindeberg-Levy theorem.

Example 2. The random variable $X_n (n = 1,2, \dots \dots)$ are independent and each of them has the Poisson distribution given by $P(X_n = r) = \frac{2^r}{r!} e^{-2} (r = 0,1,2, \dots \dots)$. Find the probability that the sum $Y_{100} = X_1 + X_2 + \cdots .. + X_{100}$ is greater than 190 and less than 210.

Solution.

$$Y_{100} = X_1 + X_2 + \cdots \dots + X_{100}$$

The random variable Y_{100} has approximately the normal distribution $N(200, 10\sqrt{2})$. $\left(\because Y_n \sim N(mn; \sigma\sqrt{n})\right)$

Since ,each of the random variable X_n has $\sigma = \sqrt{2}$ and expected value $m = 2$.

We know that, $P(y_1 < Y_n < y_2) \cong \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} e^{-\frac{z^2}{2}} dz$

Where, $y_1 = z_1 \sigma\sqrt{n} + mn,$ $\quad y_2 = z_2 \sigma\sqrt{n} + nm$

Here, $y_1 = 190$ and $y_2 = 210$ and $n = 100$

Now,

$$190 = z_1 \sigma\sqrt{n} + mn$$

$$z_1 = \frac{190 - mn}{\sigma\sqrt{n}}$$

$$= \frac{190 - 200}{\sqrt{2} \times 10}$$

$$z_2 = \frac{-10}{\sqrt{2} \times 10}$$

$$z_2 = \frac{-1}{\sqrt{2}}$$

$$z_2 = -0.707$$

and

$210 = z_2 \sigma \sqrt{n} + mn$

$$z_2 = \frac{210 - mn}{\sigma\sqrt{n}}$$

$$= \frac{210 - 200}{\sqrt{2} \times 10}$$

$$= \frac{10}{\sqrt{2} \times 10}$$

$$z_2 = \frac{1}{\sqrt{2}}$$

$z_2 = 0.707$

$$\therefore P(190 < Y_{100} < 210) = P\left(-0.707 < \frac{Y_{100} - 200}{10\sqrt{2}} < 0.707\right)$$

$$= \emptyset(0.707) - \emptyset(-0.707)$$

$$= \emptyset(0.707) - \left(1 - \emptyset(0.707)\right)$$

$$= 2\emptyset(0.707) - 1$$

$$= 2 \times 0.758036 - 1$$

$$= 1.516 - 1$$

$$= 0.516$$

$\therefore P(190 < Y_{100} < 210) \cong 0.52.$

Theorem 5.5.

Suppose that the random variable $X_1, X_2, \dots \dots$ are independent and the same distribution with standard deviation $\sigma \neq 0$. Let the random variable U_n defined by

$$U_n = \frac{X_1 + X_2 + \cdots \ldots + X_n}{n}$$

Furthermore, let $F_n(v)$ be the distribution function of random variable V_n

Defined as

$$V_n = \frac{U_n - E(U_n)}{\sqrt{D^2(U_n)}}$$

Then the sequence $\{F_n(v)\}$ satisfies the relation

$$lim_{n\to\infty} F_n(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-\frac{v^2}{2}}\, dv$$

Proof.

Here, $E(U_n) = m$ and $D^2(U_n) = \frac{\sigma^2}{n}$

$$V_n = \frac{U_n - E(U_n)}{\sqrt{D^2(U_n)}}$$

$$= \frac{\frac{1}{n}\sum_{k=1}^{n} X_k - m}{\frac{\sigma}{\sqrt{n}}} t$$

$$= \frac{\sum_{k=1}^{n} X_k - mn}{\sigma\sqrt{n}}$$

$$= z_n$$

where the random variable $z_n = \frac{\sum_{k=1}^{n} X_k - mn}{\sigma\sqrt{n}}$ by theorem Lindeberg-Levy,

The sequence $\{F_n(z)\}$ satisfies relation $lim_{n\to\infty} F_n(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{z^2}{2}}\, dz$

$\therefore$ The sequence $\{F_n(v)\}$ satisfies $lim_{n\to\infty} F_n(v) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{v} e^{-\frac{v^2}{2}}\, dv$

Hence the proof.

Remark.

Let v_1 and v_2 be two arbitrary number with $v_1 < v_2$.

$$lim_{n\to\infty} p(v_1 < V_n < v_2) = \frac{1}{\sqrt{2\pi}} \int_{v_1}^{v_2} e^{-\frac{v^2}{2}}\, dv$$

Let $u_1 = \frac{v_1\sigma}{\sqrt{n}} + m; \quad u_2 = \frac{v_2\sigma}{\sqrt{n}} + m \quad \to (1)$

$P(u_1 < U_n < u_2) \cong \frac{1}{\sqrt{2\pi}} \int_{v_1}^{v_2} e^{-\frac{v^2}{2}}\, dv$, where v_1 and v_2 are determined from (1)

∴ The random variable U_n has an asymptotically normal distribution $N\left(m; \frac{\sigma}{\sqrt{n}}\right)$.

Example 3.

The random variables $X_1, X_2, \ldots \ldots$ are independent and have the uniform distribution defined by

$$f(x) = \begin{cases} 1 & for\ x\ in\ the\ interval\ [0,1] \\ 0 & for\ \ x < 0 \quad x > 1 \end{cases}$$

Consider the random variable $Y_n = \frac{X_1 + X_2 + \cdots .. X_n}{n}$, for n=48 compute the probability than Y_n will be smaller than 0.4.

Solution.

$$f(x) = \begin{cases} 1 & x \in [0,1] \\ 0 & for\ x < 0 \quad x > 1 \end{cases}$$

Clearly, $m = \frac{1}{2}, \quad \sigma = \frac{1}{\sqrt{12}}$ (find using $E(X)$ formula)

By theorem 6.8.2, $Y_n \sim N\left(m; \frac{\sigma}{\sqrt{n}}\right)$

Here, $n = 48$

To find $P(Y_n < 0.4)$

$$P(Y_n < 0.4) = p\left(\frac{Y_n - \frac{1}{2}}{\frac{1}{\sqrt{576}}} < \frac{0.4 - \frac{1}{2}}{\frac{1}{\sqrt{576}}}\right)$$

$$= p\left(\frac{Y_n - \frac{1}{2}}{\frac{1}{24}} < -2.4\right)$$

$$\cong \emptyset(-2.4)$$

$$\cong 0.0082$$

5.6. Poisson's Chebysheve's and Khintchin's Laws of Large Numbers

Consider first a sequence of random variables $\{X_k\}$ $(k = 1,2, \dots.)$; the only assumption we make is that for every k first two moments exist, that is,

$$E(X_k) = m_k, \qquad E[(X_k - m_k)^2] = \sigma_k^2$$

Theorem 5.6. (Chebyshev's Theorem)

Let $\{X_k\}$ $(k = 1,2, \dots \dots)$ be an arbitrary sequence of random variables with variances σ_k^2. If the Markov condition $lim_{k\to\infty}\sigma_k^2 = 0$ is satisfied, the sequence $\{X_k - m_k\}$ is stochastically convergent to zero.

Proof.

Chebyshev's inequality, we have or every k and $\varepsilon > 0$

$$P(|X_k - m_k| \geq \varepsilon) \leq \frac{\sigma_k^2}{\varepsilon^2} \text{......(1)}$$

If the Markov condition

$$lim_{k\to\infty}\sigma_k^2 = 0 \dots \dots (2)$$

is satisfied, from formula (1) we obtain

$lim_{k\to\infty} P(|X_k - m_k| \geq \varepsilon) = 0.$

Thus, the sequence $\{X_k - m_k\}$ is stochastically convergent to zero.

Corollary

Let $\{X_k\}\,(k = 1,2,\ldots..)$ be a sequence of random variables pairwise uncorrelated and let $E(X_k) = m_k$ and $D^2(X_k) = \sigma_k^2$. If condition $im_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 = 0$ is satisfied, then the sequence

$$\left\{Y_n - \frac{m_1 + m_2 + \cdots + m_n}{n}\right\} \qquad (n = 1,2,\ldots..)$$

is stochastically convergent to 0.

Proof.

Suppose that the X_k considered in the last theorem are pairwise uncorrelated. Consider the random variable

$$Y_n = \frac{X_1 + X_2 + \cdots. + X_n}{n}$$

We have

$$E(Y_n) = \frac{1}{n} \sum_{k=1}^{n} m_k$$

Since the X_k are pairwise uncorrelated, we have

$$D^2(Y_n) = \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2$$

If

$$lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 = 0,$$

Then by the Chebyshev theorem, it follows that

$$lim_{n\to\infty} P[|Y_n - E(Y_n)| \geq \varepsilon] = 0$$

Thus, then the sequence

$$\left\{Y_n - \frac{m_1 + m_2 + \cdots + m_n}{n}\right\} \qquad (n = 1,2, \ldots \ldots)$$

is stochastically convergent to 0.

Remark

We considered the Poisson scheme and the generalized binomial distribution associated with it. In this scheme we consider the sum of n independent random variables $X_k \ (k = 1,2, \ldots \ldots, n)$ with the zero-one distribution, where $P(X_k = 0) = 1 - p_k, P(X_k = 1) = p_k$. Since $D^2(X_k) = p_k(1 - p_k) \leq \frac{1}{4}$, condition $lim_{n\to\infty} \frac{1}{n^2} \sum_{k=1}^{n} \sigma_k^2 = 0$ is satisfied. Thus the corollary of the Chebyshev theorem takes a form which could be called the Poisson law of large numbers.

Theorem 5.7.

If the random variable Y_n is the arithmetic mean of the random variables X_k in the Poisson scheme,

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n},$$

then the sequence

$$\left\{Y_n - \frac{p_1 + p_2 + \cdots + p_n}{n}\right\} \qquad (n = 1,2, \ldots \ldots)]$$

is stochastically convergent to 0.

Theorem 5.8.(Chebyshev law of large numbers)

Let $\{X_k\}$ $(k = 1,2, \dots \dots)$ be a sequence of pairwise un-correlated random variables with the same expected value and the same standard deviation, and let Y_n be given by formula $Y_n = \frac{X_1+X_2+\cdots+X_n}{n}$. Then the sequence $\{Y_n\}$ is stochastically convergent to the common expected value m of the random variables X_k.

Proof.

Let us now consider the case where the pairwise uncorrelated random variables $X_k(k = 1,2, \dots\dots)$ have the same expected value and the same standard deviation. Thus, for every k we can write

$$E(X_k) = m, \quad D^2(X_k) = \sigma^2$$

If we introduce the random variables Y_n defined by $Y_n = \frac{X_1+X_2+\cdots+X_n}{n}$ we have

$$E(Y_k) = m, \quad D^2(Y_n) = \sigma^2/n$$

Thus

$$lim_{n\to\infty} D^2(Y_n) = 0$$

According to the corollary of the Chebyshev theorem, the sequence $\{Y_n - m\}$ is stochastically convergent to zero.

Theorem 5.9.(Khintchin's law of large numbers)

Let $\{X_k\}$ $(k = 1,2, \dots\dots)$ be a sequence of independent random variables with the same distribution and wit expected value $E(X_k) = m$. Then the sequence $\{Y_n\}$, where

$$Y_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

is stochastically convergent to m.

Proof.

Let $\emptyset(t)$ be the common characteristic function of the random variables X_k.

By the independence of the X_k, the characteristic function of Y_n is

$$\left[\emptyset\left(\frac{t}{n}\right)\right]^n \text{....(1)}$$

Since the expected value m exists, we can expand $\emptyset(t)$ in the neighborhood of $t = 0$ according to the MacLaurin formula,

$$\emptyset(t) = 1 + mit + o(t) \dots\dots (2)$$

Substituting the expression (2) into (1), we obtain

$$\left[\emptyset\left(\frac{t}{n}\right)\right]^n = \left[1 + \frac{mit}{n} + o\left(\frac{t}{n}\right)\right]^n$$

Proceeding as in the proof of the de Moivre-Laplace theorem, we obtain

$$lim_{n\to\infty}\left[\emptyset\left(\frac{t}{n}\right)\right]^n = e^{mit} \quad \dots\dots (3)$$

The right-hand side of formula (3) is the characteristic function of the random variable Y with the one-point distribution such that

$$P(Y = m) = 1$$

By the Levy-Cramer theorem, the sequence $\{F_n(y)\}$ of distribution functions of Y_n converges to the distribution function of the random variable Y.

Thus, by theorem 6.2.1, the sequence $\{Y_n\}$ is stochastically convergent to m.

Course Material Prepared by

C. SUNDARARAJAN

Assistant Professor,

Department of Mathematics,

Manonmaniam Sundaranar University College,

Cheranmahadevi. Tirunelveli.