I M.Sc., Maths - FUZZY MATHEMATICS AND STATISTICS

SYLLABUS

Unit 1	:	The concept of Fuzziness - Some algebra of Fuzzy sets
Unit II	:	Fuzzy quantities - Logical aspects of fuzzy sets
Unit III	:	Distribution of Random variables
Unit IV	:	Conditional Probability and Stochastic independence - Some special distributions
Unit V	:	Distributions of Functions of random variables - Limiting distributions.

TEXTS :

(1) H. T. Nguyen and E.A. Walker,
A first course in Fuzzy logic (Second Edition) CRC (Chapter 1 to 4)

(2) R.V. Hagg and A.T. Craig, Introduction to mathematical statistics (Fourth Edition) Macmillan

INDEX

UNIT -1

THE CONCEPT OF FUZZINESS - SOME ALGEBRA OF FUZZY SETS

TABLE OF CONTENTS

1.1 BASIC CONCEPTS OF FUZZY SETS

This section introduces some of the basic concepts and terminology of fuzzy sets. To illustrate some of the concepts, we consider the membership grades of the elements of a small universal set in four different fuzzy sets as listed in Table 1.2 and graphically expressed in fig. 1.1 Here the crisp universal set X of ages that we have selected is table 1.2

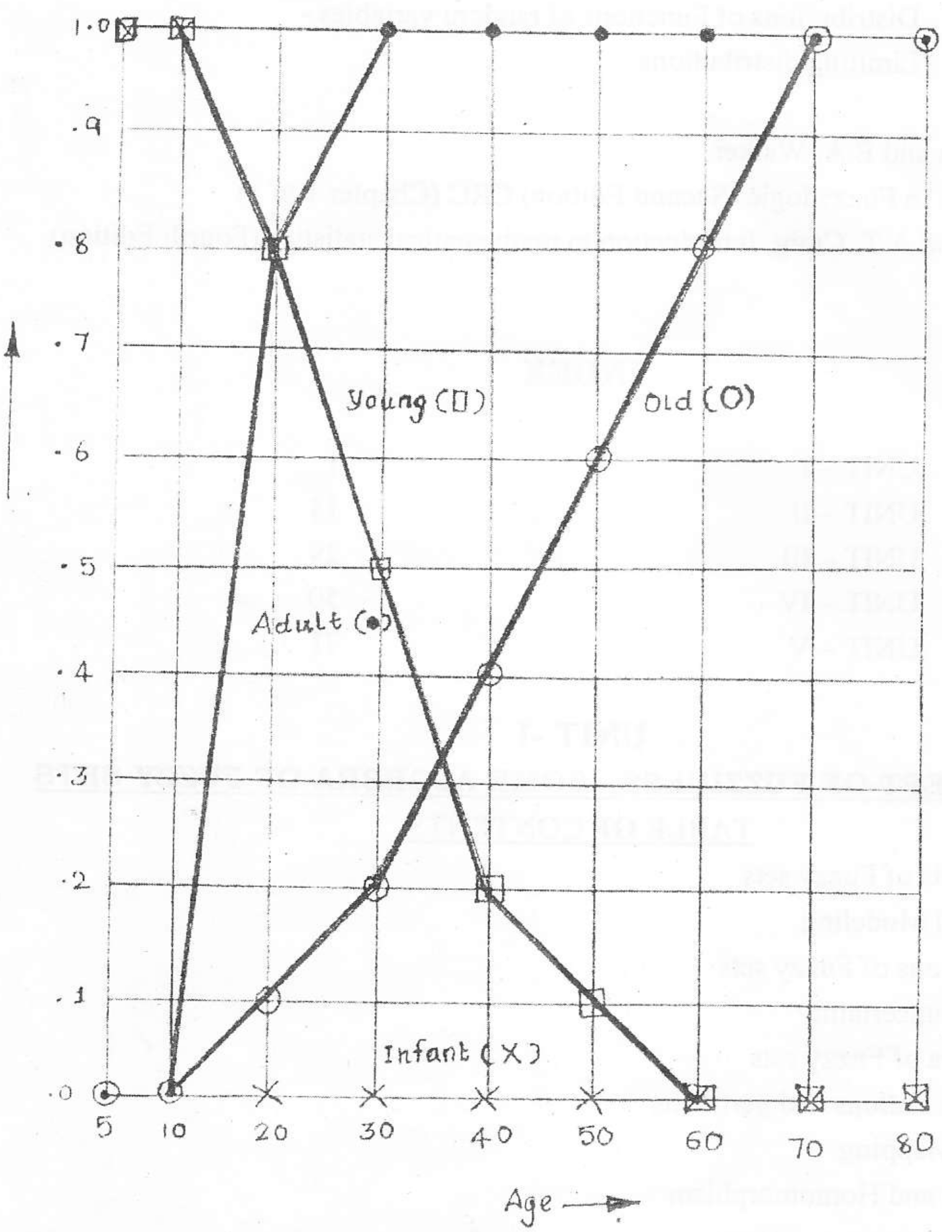


Table 1.2 Examples of Fuzzy Sets

Elements (ages)	Infant	Adult	Young	Old
5	0	0	1	0
10	0	0	1	0
20	0	.8	.8	.1
30	0	1	.5	.2
40	0	1	.2	.4
50	0	1	.1	.6
60	0	1	0	.8
70	0	1	0	1
80	0	1	0	1

If the membership grade of each element of the universal set X in fuzzy set A is less than or equal its membership grade in fuzzy set B. Thus, if

$$\mu_A(x) \le \mu_B(x),$$

for every $x \in X$, then

$$A \subseteq B.$$

The fuzzy set old from Table 1.2 is a subset of the fuzzy set adult since for each element in our universal set

$$\mu_{old}(x) \le \mu_{adult}(x).$$

Fuzzy sets A and B are called equal if $\mu_A(x) = \mu_B(x)$ for every element $x \in X$. This is denoted by

$$A = B.$$

Clearly, if A=B, then $A \subseteq B$ and $B \subseteq A$.

If fuzzy sets A and B are not equal ($\mu_A(x) \ne \mu_B(x)$ for at least one $x \in X$), we write

$$A \ne B.$$

None of the four fuzzy sets defined in Table 1.2 is equal to any of the others.

Fuzzy set A is called a proper subset of fuzzy set B when A is a subset of B and the two sets are not equal; that is, $\mu_A(x) \le \mu_B(x)$ for every $x \in X$ and $\mu_A(x) < \mu_B(x)$ for at least one $x \in X$. We can denote this by writing

$$A \subset B \text{ if and only if } A \subseteq B \text{ and } A \ne B.$$

It was mentioned that the fuzzy set old from Table 1.2 is a subset of the fuzzy set adult and that these two fuzzy sets are not equal. Therefore, old can be said to be a proper subset of adult.

When membership grades range in the closed interval between 0 and 1, we denote the complement of a fuzzy set with respect to the universal set X by A and define it by

$\mu_{\bar{A}}(x) = 1 - \mu_A(x)$,

for every $x \in X$. Thus, if an element has a membership grade of .8 in a fuzzy set A, its membership grade in the complement of A will be .2. For instance, taking the complement of the fuzzy set old from Table 1.2 produces the fuzzy set not old defined as

not old = 1/5+1/10+.9/20+.8/30+.6/40+.4/50+.2/60.

X = {5,10,20,30,40,50,60,70,80},

and the fuzzy sets labeled as infant, adults young and old are four of the elements of the power set containing all possible fuzzy subsets of X, which is denoted by P(X).

The support of a fuzzy set A in the universal set X is the crisp set they contains all the elements of X that have a nonzero membership grade in A. That is supports of fuzzy sets in X are obtained by the function

Supp : P(X) ->P(X),

where

supp A = $\{x \in X \mid \mu_A(x) > 0\}$

For istance, the support of the fuzzy set young from Table 1.2 is the crisp set

supp (young) = {5,10,20,30,40,50}

An empty fuzzy set had an empty support; that is, the membership function assigns 0 to all elements of the universal set. The fuzzy set infants as defined in Table 1.2 is one example of an empty fuzzy set within the chosen universe.

Let us introduce a special notation that is often used in the literature for defining fuzzy sets with a finite support. Assume that x_i is an element of the support of fuzzy set A and that μ_i is its grade of membership in A. Then A is written as

$$A = \mu_1/x_1 + \mu_2/x_2 + \ldots\ldots + \mu_n/x_n,$$

where the slash is employed to link the elements of the support with their grades of membership A and the plus sign indicates, rather than any sort of the algebraic addition, that the listed pairs of elements and membership grades collectively form the definition of the set A. For the case in which a fuzzy set A is defined on a universal set that is finite and coutable, we may write

$$A = \sum_{i=1}^{n} \mu_i / x_i.$$

Similarly, when X is an interval of real numbers , a fuzzy set A is often written in the form

$$A = \int_X \mu_A(x)/x$$

The height of a fuzzy set is the largest membership grade attained by any element in that set. A fuzzy set is called normalized when at least one of its elements attains the maximum possible membership grade. If membership grades range in the closed interval between 0 and 1, for instance, then at least one element must have a membership grade of 1 for the fuzzy set to be considered normalized.

1.2 MATHEMATICAL MODELING

The mathematical modeling of fuzzy concepts was presented by Zadeh in 1965. His contention is that meaning in natural language is a matter of degree. If we have a proposition such as "John is young", then it is not always possible to assert that it is either true or false. When we know that John's age is x, then the "truth", or more correctly, the "compatibility" of x with "is young", is a matter of degree. It depends on our understanding of the concept "young". If the proposition is "John is under 22 years old" and we know john's age ,then we can give a yes or no answer to whether the proposition is true or a bit by considering possible ages to be the interval $(0,\infty)$,letting A be the subset $\{x:x\in(0,\infty):x<20\}$,and then determining whether or not John's age is in A. But "young" cannot be defined as an ordinary subset of$(0,\infty)$.zadeh was led to the notion of a fuzzy subset. Clearly ,18 and 20 year olds are young, but with different degrees: 18 is younger than 20. This suggests that membership in a fuzzy subset should not be on a 0 or 1 basis, but rather on a 0 to1 scale, that is, the membership should be an element of the interval [0 ,1]. This is handled as follows. An ordinary subset A of a set U is determined by its indicator function, or characteristic function x_A defined by

$$x_A(x) = \begin{cases} 1 \text{ if } x\in A \\ 0 \text{ if } x\notin A \end{cases}$$

The indicator function of a subset A of a set U specifies whether or not an element is in A. It either is or is not. There are only two possible values the indicator function can take. This notion is generalized by allowing images of elements to be in the interval [0,1]. rather than being restricted to the two elements set {0,1}.

Definition : *A fuzzy subset of a set U is a functions* $U\rightarrow[0,\infty]$.

Those functions whose images are contained in the two element set {0,1} correspond to ordinary, or crisp subsets of U, so ordinary subsets are special cases of fuzzy subsets. A specific function $U\rightarrow[0,1]$ representing this notion would be denoted μ_A.

For a fuzzy set A: $U\rightarrow[0,1]$,the function A is called the membership function, and the value $A(\mu)$ is called the degree of membership of μ in the fuzzy set A. It is not meant to convey the likelihood or probability that μ has some particular attribute

Of course, for a fuzzy concept, different functions, A can be considered. The choice of the function A is subjective and context dependent and can be a delicate one. But the flexibility in the choice of A is useful applications, in fuzzy control.

Here are two examples of how one might model the fuzzy concept "young". Let the set of all possible ages of people be the positive real numbers. One such model, decided upon by a teenager might be

Y(x) ={ 1 if x<25
{ 40-x/15 if 25<=x<= 40
{ 0 if 40<x

1.3 SOME OPERATIONS ON FUZZY SETS

A subset A of a set U can be represented by a function x_A:U ->{0,1}, and a fuzzy subset of U has been defined to be a function A:U->{0,1}. On the set p(U) of all subsets of U there are the familiar operations of union, intercection, and complement. These are given by the rules

AUB= {x: x∈A or x∈B}
A∩B= {x: x∈A and x∈B}
A'= {x∈U: x∈A}

Operations between fuzzy sets : Consider the two fuzzy sets A(x) and B(x) of the nonnegative real numbers by the formulas

A (x) = { 1 if x<20
{ 40-x/20 if 20≤ x<40
{ 0 if 40≤x

and

$$B(x) = \begin{cases} 1 & \text{if } x\leq 25 \\ \dfrac{(1+((x-25))^2)^{-1}}{5} & \text{if } 25\leq x \end{cases}$$

Here are the plots of these two membership functions.

1.4 FUZZINESS AS UNCERTAINTY

Fuzzy sets deal with the type of uncertainty that arises when the boundaries of a class of objects are not sharply defined. The modeling of fuzzy concepts by fuzzy sets leads to the possibility of giving mathematical meaning to natural language statements. For example , when modeling the concept "young" as a fuzzy subset of [0,∞] with a membership function A: [0,∞)→[0,1], we described the meaning of "young" in a mathematical way. It is a function, and can be manipulated mathematiclly and combined with other functions.

There is a more formal relation between randomness and fuzziness. Let A :U→[0,1] be a fuzzy set. For α∈[0,1], let A_α={u∈U:A(u)≥α}. The set A_α is called the α-cut of A. Now let us view α as a random variable uniformly distributed on [0,1]. That is, let (Ω,A,P) be a probability space and α : Ω→R a random variable with

$P\{\omega : \alpha(\omega) \leq a\} = \{$ 0 if a<0

A if $0 \leq a \leq 1$

1 if a>1

Then $A_\alpha(\omega)$ is a random set.

Example :- Suppose that the illness under consideration is manifested as subsets of the set $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ of possible symptoms. Let U be a set of humans, and let S : $\rightarrow$ p(U) be given by $S(\omega) = \{u \in \Omega$:u has symptom $\omega\}$. For $u \in U$, we are interested in some numerical measure of the set $\{\omega \in \Omega : u \in S(\omega)\}$. This is to be a measure of the seriousness of the illness of u. Medical experts often can provide assessments which can be described mathematically as a function u: $P(\Omega) \rightarrow [0,1]$, where $\mu(B)$ is the degree of seriousness of the illness of a person having all the symptoms in B. So a membership function can be taken to be

$$A(u) = \mu\{\omega \in \Omega : u \in s(\omega)\}$$

Since μ is subjective, there is no compelling reason to assume that it is a measure.

1.5 SOME ALGEBRA OF FUZZY SETS

1.5.1 Boolean algebras and lattices

Definition : A relation on a set U is a subset R of the cartesian product U ×U .

The notion of relation is very general one. For an element(x, y) $\in U \times U$ either(x, y)$\in$R or it is not.

The relation $\subseteq$ satisfies the following properties.

$A \subseteq A$(the relation reflexive)

If $A \subseteq B$ and $B \subseteq C$ then $A \subseteq C$.(the relation is transitive)

If $A \subseteq B$ and $B \subseteq A$, then A = B.(the relation is antisymmetric)

A partial order on a set is a relation on that set that is reflexive ,transitive and antisymmetric.

Definition: A partially ordered set is a pair(U,$\leq$)where U is a set and $\leq$ is a partial order on U.

Definition: A lattice is a partially ordered set (U,$\leq$)in which every pair of elements of U has a sup and an inf in U.

Chains are always lattices. The partially ordered set (P(U), $\subseteq$)is a lattice. The sup of two elements in P(U) is their union, and the inf is their intersection .The interval[0,1]is a lattice, being a chain.

Lemma: 1.5.2 If (U, $\leq$)is a lattice, then for all a, b, c$\in$ U,

1. $a \vee a = a$ and $a \wedge a = a$ (V and $\wedge$ are indempotent.
2. $a \vee b = b \vee a$ and $a \wedge b = b \wedge a$(V and $\wedge$ are commutative)

3. (aVb)Vc=aV(bVc) and (a Λ b)Λc= aΛ(bΛc). (V and are associative).

4. aV (a Λ b) = a and aΛ(aVb) =a. These are the absorption identities).

Theorem : .1.5.3 If U is a set with binary operations V and Λ which satisfy the properties of Lemma 1.5.2, then defining a ≤ b if a∧b = a makes (U, ≤) into a lattice whose sup and in operations are V andΛ.

Proof. We first show that a ∧ b =a if and only if a V b=b. Thus defining a≤b if aΛb=a is equivalent to defining a≤b if a V b=b. Indeed, if aΛb=a, then aVb=(aΛb) V b=b by one of the absorption laws. Similarly, if aVb=b, then a∧b=a.We show the existence of sups, and claim that sup{a,b}=aVb. Now a≤aVb since a∧ (aVb)=a by one of the absorption laws. Similarly b≤bVa=aVb, so that aVb is an upper bound of a and b. For any other upper bound x,a=a∧x and b=bΛx, whence x = aVx = bVx. Therefore, x=aVxΛb a=(aVb)V x, and so aVb≤x. Thus aΛb=sup{a,b}, Hence, the proof follows.

The lattice ([0,1],≤) plays a fundamental role. It is a bounded distributive lattice. It is not complemented. For x,y∈[0.1], xVy = sup{x,y}= max{x,y}, and similarly x Λ y=inf{x,y}. Distributivity is easy to check. This lattice has another important operation on it. [0,1]→[0,1]:x→1-x. We denote this operation by ' even though it is not a complement.The operation has the following properties

(x')' = x

x≤y implies that y'≤x'

such an operation on a bounded lattice is called an involution, or a duality. It follows that ' is one-to-one and onto, and that 0'=1 and 1'=0. If is an involution, the equations

$(x \vee y)^1 = x^1 \wedge y^1$

$(x \wedge y)^1 = x^1 \vee y^1$

are called the De Morgan laws.

Theorem : 1.5.4

Let (V,V, Λ ,',0,1) be a De Morgan algebra and let ∪ be any set. Let *f* and g be mappings from U and V. We define

1. *f* V g)(x)= *f* (x) Vg(x)

2.(*f*Λ g)(x)= *f* (x) Λ g(x)

3. *f* '(x)=(*f* (x)) ^{1}

4.0(x)=0

5.1(x)=1

let V^U be the set of all mappings from U into V. Then $(V^U, \wedge, ', 0, 1)$ is a De Morgan algebra. If V is a complete lattice, then so is V^U

Proof :

The proof is routine in all respects. For example, the fact that V is an associative operation on V^U comes directly from the fact that V is associate on v. (the two Vs are different of course) Using the definition of v^U and that v is associate on v ,we get

$$\begin{aligned}(f \vee (g \vee h))(x) &= f(x) \vee (g \vee h)(x)\\ &= f(x) \vee (g(x) \vee h(x))\\ &= (f(x) \vee g(x)) \vee h(x)\\ &= (f \vee g)(x) \vee h(x)\\ &= ((f \vee g) \vee h)(x).\end{aligned}$$

Whence $f \vee (g \vee h) = (f \vee g) \vee h$ and so V is associative on V^U and hence directly proof follows.

Corollary 1.5.5

$f(u), \vee, \wedge, ', 0, 1)$ is a complete De Morgan algebra.

1.6 EQUIVALENCE RELATIONS AND PARTITIONS

Definition : A relation ~on a set U is an equivalence relation if for all a, b and c in U.

(1) a~a

(2) a~b implies b~a, and

(3) a~b,b~c imply that a~c.

The first and third conditions we recognize as reflexivity and transitivity. The second is that of symmetry. Thus an equivalence relation is a relation that is reflexive, symmertric and transitive.

Definition: Let ~ be an equivalence relation on a set U and let a∈U.The equivalence class of an element a is the set [a]={u∈U:u~a}.

Definition: Let U be a nonempty set.A partition of U is a set of nonempty pairwise disjoint subsets of U whose union is U.

Theorem: 1.6.1

Let ~ be an equivalence relation on the set U.Then the set of equivalence classes of ~ is a partition of U.This association of an equivalence relation ~ with the partition consisting of the equivalence classes of ~ is a one-to-one correspondence between the set of equivalence relations on U and the set of partitions of U.

Proof : The union of the equivalence classes [u] is U since u ι[u].We need only that two equivalence classes be equal or disjoint. If x ε [U] ⊓[V] {v},then x~u,x~u and so u~x and x~v.By transitivity that u ~ v. If y ε [u], then y ~ v and so u~x and x~ v.It follows y~v.Thus y ∈[v].This means that {u} ⊆ [V]. Similarly, [v]⊆[u] and hence [u] = [v]. So if two equivalence classes are not disjoint they are equal. Therefore the equivalence classes from a partition. Notice that two elements are equivalence are equivalent if and only if they are in the same member of the partition, that is in the same equivalence class. So this map from equivalence relations to partitions that one-to-one. Given a partition declaring two elements equivalent if they are in the same member of the partition that is, in the same equivalence class. So the map from equivalence relations to partitions is onto.

Theorem : *Let $\epsilon(U)$ be the set of all equivalence relations on the set U. Then $(\epsilon(U), \subseteq)$ is a complete lattice.*

Proof: There is a biggest and smallest element of $\epsilon(U)$, namely U□U and $\{(u,u): u \in U\}$, respectively. We have to show that any nonempty family $\{E_i : I \in I\}$ of elements of $\epsilon(U)$ has a sup and inf. Now certainly $\Lambda \{E_i : i \in I\} = \Omega\, I\, E_i$ if $\cap_{i\in I} E_i$ is an equvalence relation. Let (u,v) and $(v,w) \in \cap_{i\in I} E_i$. Then (u,v) and (v,w) belong to each E_i and hence (u,w) belongs to each E_i. Therefore, $(v,w) \in \cap_{i\in I} E_i$. Thus $\cap_{i\in I}$ is a transitive relation on U. That $\cap_{i\in I} E_i$ is reflexive and symmetric is similar. What we have shown is that the intersection of any family of equivalence relations on a set is an equivalence relation on that set. This is clearly the inf of that family. Now $V\{E_i : i \in I\}$ of a family of equivalence relations on U is

$$\cap\{E \in (U) : E \supseteq E_i \text{ for all } i \in I\}$$

Note that U×U is an equivalence containing all the Ei. This intersection is an equivalence relation on U and it is clearly the least equivalence relation containing all the Ei. Therefore it is the desired sup.

1.7 COMPOSING MAPPING

Let $f : U \rightarrow V$, and $g : V \rightarrow W$. Then $g \circ f$, or more simply gf, is the mapping $U \rightarrow W$ defined by $(gf)(u) = g(f(u))$. This is called the **composition** of the mappings f and g. Any two functions of a set into itself can be composed. The function $f : U \rightarrow U$ such that $f(u)=u$ for all u is denoted by IU and is called the **identity function** on U. The set of all functions from U to V is denoted map(U,V),or by V^U.

A mapping $A: U \rightarrow L$ induces a mapping $A : p(u) \rightarrow P(L)$. So with a sub set X of U. A(X) is a subset of L. But since L is a complete lattice. We may take the sup of A(X). This sup is denoted V(A(X). One should view V as a mapping $p(L) \rightarrow L$. The composition V A is a mapping $P(U) \rightarrow L$,namely the mapping given by

$$P(U) \xrightarrow{A} P(L) \xrightarrow{V} L$$

In particular, *a fuzzy subset of U yields a fuzzy subset of P(U).*

For sets U and V, a subset of U×V is called a relation in U×V. Now a relation R in U×V induces a mapping $R^{-1}: V \rightarrow P(U)$ given by

$$R^{-1}(v) = \{u : (u, v) \in R\}$$

Thus with A: U→L we have the mapping

$V \xrightarrow{R^{-1}} P(U) \xrightarrow{A} P(L) \xrightarrow{V} L$

Thus the relation R in U×V associates with a mapping A: U→L a mapping V AR^{-1}: V→ L. This mapping is sometimes denoted R(A). When L=[0,1], we then have a mapping F(U)→F(V) sending A to R(A)=VAR^{-1}. If R is actually a function from U to V, then R has been *extended* to a function F(U)→ F(V) sending A to VAR^{-1}. In fuzzy set theory, this is called **extension principle.**

1.8 ISOMORPHISMS AND HOMOMORPHISMS

The mapping f(x) = x+1 is an order isomorphism from [0,1]to [1,2]. A mapping g : U→V such that g(x)≤ g(y) whenever x≤y is called **homomorphism**, or an **order homomorphism**, emphasizing that the order relation is being respected. The condition on g that if x ≤y then g(x)≤g(y) is expressed by saying that g preserves order or is **order preserving.**

A mapping f : U→V is an **isomorphism** of two lattices if f is one- to -one and onto, f(x V y)=f(x) V f(y) and f(x∧y)= f(x)∧f(y). That is, f must be one-to-one and onto and preserve both lattice operations. If the one-to-one and onto conditions are dropped, then f is a **lattice homomorphism**. If U and V are complete lattices, then an isomorphism f : U→V is a complete **lattice homomorphism** if and only if f (VS) = V{f(s):S ∈S} and f(∧s)=∧{f(s):s ∈S} for every subset S of U. An isomorphism of a lattice(or any algebraic structure) with itself is called an **automorphism.**

Example :

Consider the lattice([0,1],V,∧,')with involution, where V is sup, ∧ is inf, and x'=1-x,and the lattice{0,1/2,1}with the same operations, Then the mapping f:[0,1]→{0,1/2,1} that sends endpoints to endpoints and the interior points of[0,1]to ½ is a homomorphism. Note that one requirement is that f(x')=f(x)',and that this does hold.

Suppose that f: U→V is a homomorphism from a lattice(U,V,∧) to a lattice(V,v,∧).Then the relation ~ on U by a ~ b if f(a)=f(b) is an equivalence relation. But also if a ~ b and c ~ d then f(aV c)=f(a) V f(c)=f(b) V f(d)=f(b V d), so a v c and c ~ d. Similarly a∧c ~ b∧d. So this equivalence relation has these two additional properties :if a ~ b and aVc ~ bVd and a∧c ~ b∧d. Such an equivalance relation on a lattice is called a **congruence**. And congruences on lattices give rise to homomorphisms.

1.9 ALPHA-CUTS

Definition :-

Let U be a set, let C be a partially orderded set and let A: U→C. For α ∈C , the α-cut of A, or the α-level set of A, is A^{-1}(↑α)={u∈U : A(u)≥ α}. This subset of U will be denoted by A_α.

Thus the α-cut of a function $A: U \to C$ is the subset $A_\alpha = A^{-1}(\uparrow\alpha)$ of U, and we have one such subset for each $\alpha \in C$. A fundamental fact about the α-cuts, A_α is that they determine . It follows immediately from the equation

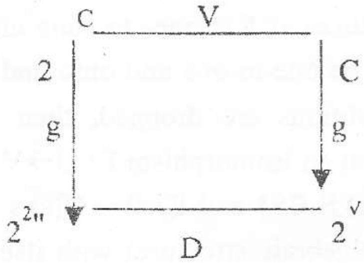
$$A^{-1}(\alpha) = A_\alpha \cap \bigcap_{B>\alpha} (\cup A_\beta)^1$$

Result :

Let A and B be a mappings from a set U into a partially orderded set C. If $A_\alpha = B_\alpha$ for all $\alpha \in C$, then A=B.

Theorem : 1.9.1

Let C be a complete lattice and U a set. Let F(U) be the set of all mappings from U into C, and L(U) be the set of all mappings g: C→P(U) such that the diagram given below commutes or equivalently such that for all subsets D of C,



$$g(VD) = \bigcap_{d\in D} g(d)$$

Then the mapping φ : F(U)→L(U) given by $\varphi(A)=A^{-1}\uparrow$ is one-to-one and onto.

Proof.

We have already observed that ϕ maps F(U) into l(u) and that this mapping is one-to-one. Let g ∈ L(U). We must show that g $=A^{-1}$ for some A∈F (U). For u ∈U, define

$$h(u) = \{d \in C: g(d) \supseteq \bigcap_{u\in g(x)} g(x)\} = \{d \in C: u \in g(d)\}$$

Let A=Voh. Then

$$A^{-1}\uparrow(c) = \{u \in U: A(u) \geq 0\}$$

Now if $u \in g(c)$, then $g(c) \supseteq \bigcap_{u\in g(x)} g(x)$ which implies that cιh(u)

and thus that is $\in A^{-1}(c)$.I y(c)$\subseteq A^{-1}$. Now suppose that uι$A^{-1}\uparrow(c)$. Now suppose that uι $A^{-1}\uparrow(c)$, so that $A(u) \geq c$. Then $u \in \bigcap_{u\iota g}(x)$ $g(x) \subseteq g(d)$ for all dεh(u).

Thus $u \in \bigcap_{d\iota A(u)} g(d) = g(A(U)) \subseteq g(c)$

It follows that $g(c)=A^{-1}\uparrow(c)$, whence $g=A^{-1}\uparrow=\phi(A)$.

Corollary : *The complete lattices F(U) and L(U) are isomorphic.*

1.10 IMAGES OF ALPHA-LEVEL SETS

Let $f : U \to V$ and let A be a fuzzy subset of U. Then $\vee Af^{-1}$ is a fuzzy subset of V by the extension principle. It is the mapping that is the composition.

$$V \to^{f^{-1}} P(U) \xrightarrow{A} \varpi([0,1])^V \to [0,1]$$

Theorem : 1.10.1

Let C be a complete lattice, U and V be sets, $A: U\to C$, and $f : U\to V$. then

1. $f(A\alpha)\subseteq(\vee Af^{-1})\alpha$ for all $\alpha \in C$.
2. $f(A\alpha)= (\vee Af^{-1})_\alpha$ for $\alpha >0$ if and only if for each member P of the partition induced by f, $\vee A(P) \geq \alpha$ implies $A(u) \geq \alpha$ for some $u \in P$.
3. $f(A_\alpha) = (\vee Af^{-1})_\alpha$ for all $\alpha > 0$ if and only if for each member P of the partition induced by f, $\vee A(P)=A(u)$ for some $u \in P$.

Proof : The theorem follows immediately from the equalities below.

$$f(A_\alpha) = \{f(u) : A(u)\geq\alpha\}$$
$$= \{v \in V : A(u) \geq \alpha, f(u)=v\}$$
$$(\vee Af^{-1})_\alpha = \{v \in V : \vee Af^{-1}(v) \geq \alpha)$$
$$= \{v \in V : \vee\{A(u) : f(u) = v\}\geq\alpha\}$$

One should notice that for some α, it may not be true that $\vee A(P) = \alpha$ for any P.

EXERCISES

1. Let U be a set and P(U) be the set of all subsets of U. Verify in detail that $(P(U),\subseteq)$ is a Boolean algebra. Show that it is complete.
2. Show that a chain with more than two elements is not complemented.
3. Show that the De Morgan algebra $(F(U),\vee,\wedge,',0,1)$ satisfies $A\wedge A' \leq B \vee B'$ for all $A,B \in F(U)$, that is, is a Kleene algebra, Show that [0,1] is a Kleene algebra. Show that $[0,1]^{[2]}$ is not a Kleene algebra.
4. Let B be a Boolean algebra . Show that $B^{[2]}$ is a Stone algebra but not a Boolean algebra.
5. Show that if S is a Stone algebra, then so is $S^{[2]}$.
6. Show that every bounded chain is a stone algebra.

UNIT - II

FUZZY QUANTITIES - LOGICAL ASPECTS OF FUZZY SETS

TABLE OF CONTENTS

2.1 Fuzzy quantities

Let R denote the set of real numbers. The elements if F(R), that is, the fuzzy subsets of R, are Fuzzy quanitities. A relation R in U x V. Which is simply a subset R of U x V. induces the mapping R: $f(U) \rightarrow f(V)$ defined by R(A)= VAR-1. This is the mapping given by

$$R(A)(v) = V\{A(\{u:u,v)\in R\})\}$$

are expressed by the extension principle at work. In particular, a mapping $f : R \rightarrow R$ induces a mapping $f: f(R) \rightarrow f(R)$. A binary operation o: $Rx R \rightarrow R$ gives a mapping $f(RxR) \rightarrow f(R)$, and we have the mapping $f(R) \times f(R) \rightarrow f(RxR)$ sending (A,B) to $\Lambda(AxB)$. Remember that Λ $(AxB)(r,s) = A(r) \wedge B(s)$. The composition

$$F(R) \times F(R) \rightarrow F(RxR) \rightarrow F(R)$$

of these two is the mapping that sends (A,B) to $V(\Lambda(AxB))o^{-1}$. Where $o^{-1}(x) = \{(a,b) : aob = x\}$. We denote this binary operation by A o B.

This means that

$$(AoB)(x) = V\Lambda(AxB)\, o^{-1}(x)$$
$$= V_{aob=x}\, \Lambda(AxB)(b)$$
$$= V_{aob=x}\, \{A(a) \wedge B(b)\}$$

For example, for the ordinary arithmetic binary operations of addition and multiplication on R, we then have corresponding operations $A+B = V\Lambda(AxB)+^{-1}$ and $A.B = V\Lambda(AxB).^{-1}$ on F(R). Thus

$$(A+B)(z) = V_{x+y=z}\, \{A(x) \wedge B(y)\}$$
$$(A.B)(z) = V_{x.y=z}\, \{A(x) \wedge B(y)\}$$

The mapping $R \to R : r \to -r$ induces a mapping $f(R) \to f(R)$ and the image of A is denoted -A

For $x \in R$,

$$(-A)(x) = V_{x=-y} \{A(y)\} = A(-x)$$

If we view - as a binary operation on R, we get

$$(A-B)(z) = V_{x-y=z} \{A(x) \wedge B(y)\}$$

It turns out that A+(-B) = A-B, as is the case for R itself.

Division deserves some special attention. It is not a binary operation on R since it is not defined for pairs (x,0), but it is the relation

$$\{((r,s),t) \in (RxR) \times R : r = st\}$$

By the extension principle, this relation induces the binary operation on f(R) given by the formula

$$A/B(x) = V_{y=zx} (A(y) \wedge B(z))$$

Proposition 1

For any fuzzy set A, A/x{0} is the constant function whose value is A(0).

Proof. The function A/x{o} is given by the formula

$$(A/x\{o\})(u) = V_{s=t.u} (A(s) \wedge x\{o\}(t))$$

$$= V_{s=o.u} (A(s) \wedge x\{o\}(0))$$

$$= A(0)$$

Theorem 2 Let o be any binary operation on a set U, and let S and T be subsets of U. Then

$$T_S{}^o T_T = T\{sot : s \in s, t \in T\}$$

Proof. For $u \in U$,

$$(T_S \circ T_T)(u) = V_{sot} = u (T_S(s) \wedge T_T(t))$$

The sup is either 0 or 1 and is 1 exactly when there is an $s \in S$ and a $t \in T$ with sot = u. The result follows.

Theorem : 3

Let A, B and C be fuzzy quantities. The following hold.

1. 0+A=A
2. 0.A=0
3. 1.A=A
4. A+B=B+A
5. A+(B+C) = (A+B)+C
6. AB=BA
7. (AB)C=A(BC)
8. r(A+B) =rA+rB
9. A(B+C) ≤ AB+AC
10. (-r)A= -(rA)
11. -(-A) = A
12. (-A)B=-(AB) = A(-B)
13. A/1=A
14. A/r = 1/r A
14. A/B = A 1/B
16. A+(-B) = A-B.

Proof. We prove some of these. For the equations

$$A(x) = V_{yz=x} X_{\{1\}} (y) \Lambda A(z)$$
$$= V_{1x=x} X\{1\} (1) \Lambda A(x)$$
$$= A(x)$$

show that 1.A=A. If (A(B+C))(x) > (AB + AC) (x), then there exist u,v,y with y(u+v) = x and such that

$$A(y) \Lambda B(u) \Lambda C(v) > A(p) \Lambda B(q) \Lambda C(k)$$

for all p,q,h,k with pq + hk = x. But this is not so for p = h = y, q = u, and n = k. Thus (A(B+C))(x) ≤ (AB+AC)(X) for all x, whence A(B+C) ≤ AB+AC.

However,

$$r(A+B) = rA+rB \text{ since}$$
$$(X\{r\}(A+B))(x) = V_{uv=x} (X\{r\} (u) \Lambda (A+B) (v))$$
$$= V_{rv=x} (X\{r\}(r) \Lambda (A+B) (v))$$
$$= V_{s+t=v\ rv=x} (A(s) \Lambda B(t))$$
$$= V_{s+t=v\ rv=x} (X\{r\}(r)A(s) \Lambda X\{r\}(r)B(t))$$
$$= (rA + rB)(x)$$

Definition A fuzzy quantity A is convex if its α-cuts are convex, that is, if its α-cuts are intervals.

Theorem (4) A fuzzy quantity A is convex if and only if A(u) ≥ A(x) Λ A(z) whenever x ≤ y ≤ z.

Proof : Let A be convex, x ≤ y ≤ z, and α = A(x) Λ A(z). Then x and z are in Aα is an interval, y is an Aα. Therefore A(y) ≥ A(x) Λ A(z).

Suppose that A(y) ≥ A(x) Λ A(z) whenever x ≤ y ≤ z. Let x < y < z with x,z ∈ Aα. Then A(y) ≥ A(x) Λ A(z) ≥ α, whenever y∈ Aα and Aα is convex.

Definition

A fuzzy quantity A is convex if its α-cuts are convex, that is , if its α-cuts are intervals.

Therorem (5)

A fuzzy quantity A is convex if and only if A(y)≥A(x) Λ A(z) whenever S ≤ S ≤ Z

Proof .

Let A be convex, x≤y≤z, and α = A(x) Λ A(z). The x and z are in A_α, and since A_α is an interval, y is in A_α. Therefore A(y) ≥ A(x) Λ A(z).

Suppose that A(y) ≥ A(x) Λ A(z) whenever x≤y≤z. Let x<y<z with x,z ∈A_α. Then A(y) ≥ A(x), whence y∈A_α and A is convex.

Theorem (6)

If A and B are convex, then so are A+B and -A.

Proof : -

We show that A+B is convex . Let x<y<z. We need that (A+B)(y) $\geq$ (A+B)$\wedge$(x)(A+B)(z) . Let X>0. There are numbers X_1,X_2,Z_1 and Z_2 with $X_1+X_2 = X$ and $Z_1+Z_2 = Z$ and satisfying

$$A(X_1)\ B(X_2) \geq (A+B)(x)-\varepsilon$$

$$A(Z_1)\ B(Z_2)\ (A+B)(Z) -\varepsilon$$

Now y = ax +(1-α) z for some α∈[0,1]. Let $x^1 = ax_1+(1-\alpha)z_1$ and $z^1=ax_2+(1-)z_2$. then $x^1+z^1=y$, x^1 lies between x_1 and z_1, and z_1 lies between x_2 and z_2. Thus we have

$$\begin{aligned}(A+B)(y) &\geq A(x^1)\,B(z^1)\\ &\geq A(x_1)\wedge A(z_1)\wedge B(x_2)\wedge B(z_2)\\ &\geq [(A+B)(x)-\varepsilon[(A+B)(z)-\varepsilon]\\ &\geq [(A+B)(x)\wedge (A+B)(z)]-\varepsilon\end{aligned}$$

It follows that A+B is convex.

A function f: R->R is upper semicontinous if {x:f(x)$\geq$ α} is closed. The following definition is consistent with this terminology.

Definition

A fuzzy quantity is upper semicontinous if its α-cuts are closed.

Theorem (7)

A fuzzy quantity semicontinous if and only whenever x ∈ R and ∈>0 there is δ >0 such that |x-y| < δ implies that A(y)<A(x)+∈

Proof

Suppose that A_α is closed for all $_\alpha$. Let x ∈ R and ε >0. If A(x)+ ε >1, then A(y)<A(x)+ ∈ for any y. If A(x) + ε$\leq$1 then for α=A(x)+ε, x∈A_α and so there is δ>0 such that (x-δ, x+δ) ∩A_α = φ. Thus A(y)<α = A(x)+ ∈ for all y with x-y<.δ

Conversely, take α∈[0,1], x ∈ A_α, and ε = $\frac{\alpha\text{-}A(x)}{2}$. There is δ>0 such that x-y < δ implies that A(y)<A(x) + $\frac{\alpha\text{-}A(x)}{2}$< α and so (x-δ, x+δ) ∩ A_α = 0 . Thus A_α is closed.

The following theorem is the crucial fact that enables us to use α-cuts in computing with fuzzy quantities.

Theorem (8)

Let O : RxR->R be a continuous binary operation on R and let A and B be fuzzy quantities with closed α-cuts and bounded supports. hen for each u∈R, (AoB)(u)=A(x)$\wedge$B(y) for some x and y with u=xoy.

Proof : By definition,

$$(AoB)(u) = \bigvee_{xoy=u} (A(x) \wedge B(y))$$

The equality certainly holds if $(AoB)(u)=0$. Suppose $\alpha = (AoB(u)>0$, and $A(x)\wedge B(y)< \alpha$ for all x any y such that there is a sequence $\{A(x_i)\ B(y_i)\}$ i=1 in the set $\{A(x)\ B(y) : xoy=u\}$ having the following properties.

1. $\{A(x_i)\ B(y_i)\}$ converges to α
2. Either $\{A(x_i)\}$ or $B(y_i)\}$ converges to α
3. Each x_i is in the support of A and each yi is in the support of B

Suppose that it is $\{A(x_i)$ that converges to α. Since the support of A is bounded, the set {xi} has a limit point x and hence a subsequence converging to x. Since the support of B is bounded, the correspondent subsequence of y_i has a limit point y and hence a subsequence converging to y. The corresponding subsequence of x_i converges to x. Thus we have a subsequence $\{\{A(xi)\wedge B(yi)\}_{i=1}$ satisfying the three properties above and with $\{x_i\}$ converging to x and $\{y_i\}$ converging to y. If $A(x) = \lambda < \alpha$, then for $\delta = \frac{\alpha+\lambda}{2}$ and for sufficiently large i,

$x_i \in A_i$, x is a limit point of those x_i, and since all cuts are closed, $x\in A_i$. But it is not, so A(x) = α. In a similar vein, $B(y) \geq \alpha$ and we have $(AoB)(u)=A(x)\wedge B(y)$. Finally , u=xoy since $u=x_ioy_i$ for all i, and o is continuous.

Corollary (9)

If A and B are fuzzy quantities with bounded support, all α-cuts are closed, and o is a continous binary operation on R, then $(AoB)_\alpha = A_\alpha oB_\alpha$.

Proof :

Applying the theorem, for $u \in (AoB)_\alpha$, $(AoB)(u) = A(\alpha)\wedge B(y)$ for some x and y with u=xoy. Thus $x\in A_\alpha$ and $y\in B$, and therefore $(AoB)_\alpha \subseteq A_\alpha oB_\alpha$. The other inclusion can be calculated easily.

Corollary (10)

If A and B are fuzzy quantities with bounded support and all α-cuts are closed, then

1. $(A+B)_\alpha = A_\alpha + B_\alpha$
2. $(A.B)_\alpha = A_\alpha . B_\alpha$
3. $(A-B)_\alpha = A_\alpha - B_\alpha$

2.2 FUZZY NUMBERS

Definition

A fuzzy number is a fuzzy quantity A that satisfies the following conditions.

1. A(x) = 1 for exactly one x.
2. The support {x: A(x)>0} of A is bounded.

3. The α cuts of A are closed intervals.

Proposition (1)

The following hold:

1. Real numbers are fuzzy numbers.
2. A fuzzy number is a convex fuzzy quantity.
3. A fuzzy number is upper semicontinuous.
4. If A is a fuzzy number with A(r) =1, then A is non-decreasing on $(-\infty,1)$ and non-increasing on $[r,\infty)$.

Proof. It should be clear that real numbers are fuzzy numbers. A fuzzy number is convex since its α-cuts are intervals, and is upper semicontinuous since its α-cuts are closed. If A is fuzzy number with A (r) = 1 and $x < y < r$, then since A is convex and $A(y) < A(r)$, we have $A(x) \leq A(y)$,so A is monotone increasing on $(-\infty,r]$. Similary, A is monotone decreasing on $[r,\infty)$.

Theorem (2)

If A and B are fuzzy number then so are A+B,A.B, and –A.

Proof.

That these fuzzy quantities have bounded support and assume the value 1 in exactly one place is easy to show. The α-cuts of A+B and A.B are closed intervals by the last Corollary of ?1. Since –A=(-1). the remaining parts follows.

Definition

A triangular fuzzy number is a fuzzy quantity A whose values are given by the formula

$$A(x) = \begin{cases} 0, & \text{if } x < a \\ \dfrac{x-a}{b-a} & \text{if } a \leq x \leq b \\ \dfrac{x-c}{b-c} & \text{if } b \leq x \leq c \\ 0 & \text{if } c < x, \end{cases} \qquad \text{for some } a \leq b \leq c.$$

Theorem (3)

For triangular numbers,

$$(a, b, c) + (d, e, f) = (a+d, b+c, c+f)$$

Proof. Using $((a, b, c) + (d, e, f))\alpha = (a, b, c)\alpha + (d, e, f)\alpha$, it follows that the support of the sum is the interval (a+d, c+f) and that 1 is assumed exactly at b+e. Suppose that $\alpha > 0$, the left endpoint of the α-cu of (a, b, c)is u and that of(d, e, f) is v. Then $a \leq u \leq b$, $d \leq V \leq e$, and

$$\alpha = \frac{u-a}{b-a} = \frac{u-d}{e-d}$$

Also, by alegbrical Principle,

$$\alpha = \frac{u+v-(a+d)}{b+e-(a+d)}$$

which shows that u+v is the left endpoint of the α-cut of(a+d, b+e,c+f).

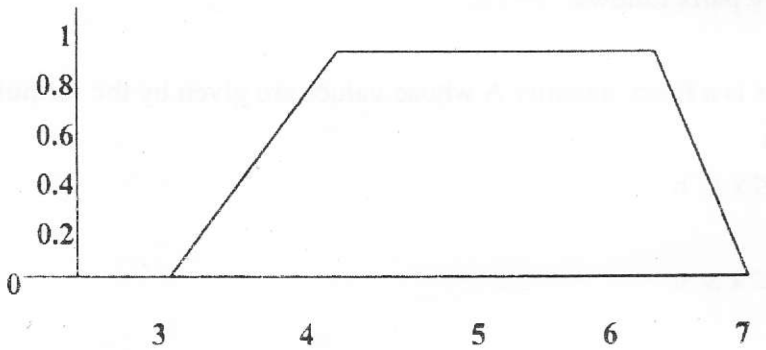
But we know that the endpoint of the α-cut of(a,b,c)+(d,e,f) is u+v. Similarly for right endpoint of cuts, and hence (a,b,c) +(d,e,f) and (a+d,b+e,c+f)have the same cuts and so equal.

2.3 FUZZY INTERVALS

A subset S of R is identified with x_s, and in particular, interval [a, b] are identified with their characteristic functions, namely the fuzzy quantities $x_{[a, b]}$.

The use of intervals with their arithmetic is appropriate in some situations involving impreciseness. When the intervals themselves are not sharply defined, we are driven to the concept of fuzzy interval. Thus we want to generalize intervals to fuzzy intervals, and certainly a fuzzy quantity generalizing the interval [a, b]. A fuzzy quantity that attains the value 1 is called normal. The other defining properties of fuzzy intervals should be like those of fuzzy numbers. Thus a fuzzy interval should look something like the following picture.

This fuzzy interval has a trapezoidal form representing "approximately between 4 and 6". Our definition is this:



Definition

A fuzzy interval is a fuzzy quantity A satisfying the following:

1. *A is normal*
2. *The support {x :A(x)>0} of A is bounded.*
3. *The α-cuts of A are closed intervals.*

2.4 LOGICAL ASPECTS OF FUZZY SETS

Any function t :V→{0,1} we get a function ~t :F→{0,1] as follows: for each variable a appearing in a formula, substitute t (a) for it. Then we have an expression in the symbols 0,1,V,∧, and ', together with balanced sets of parentheses. The tables below define the operations of V,∧ and ' on the truth values {0,1}.

V	0	1
0	0	1
1	1	1

$\wedge$	0	1
0	0	0
1	0	1

0	1
1	0

Using these tables, which describe the two element Boolean algebra, we get an extension to F. For example, if t(a)=0 and t(b)=t(c)=1, then

$$\sim t((a \vee b)\wedge c)\wedge(b' \vee c)) = (((t(a) \vee t(b)\wedge t(c))\wedge(t(b)' \vee t(c)))$$
$$= ((0 \vee 1)\wedge 1)\wedge 1' \vee 1)$$
$$= (1\wedge 1)\wedge(0 \vee 1)$$
$$= 1\wedge 1$$
$$= 1$$

Such a mapping $F\rightarrow\{0,1\}$ is called a truth evaluation. We have exactly one for each mapping $V\rightarrow\{0,1\}$. Expressions that are assigned the value 1 by every t are called tautologies. Such as $a \vee a'$ and $b \vee b'$.

There are two other common logical connectives.$\Rightarrow$(implies) and $\Leftrightarrow$(implies and is implied by, or if and only if), and we could write down the useful truth tables for them. However, in classical two-valued logic, $a\Rightarrow b$ is taken to mean $a' \vee b$, and $a\Leftrightarrow b$ to mean $(a\Rightarrow b)\wedge(b\Rightarrow a)$. Thus they can be defined in terms of three connectives we used. The formula $a\Rightarrow b$ is called material implication.

Now the set $F/\equiv$(F "modulo"$\equiv$) of all equivalence-classes of this equivalence relation. Let [a] denote the equivalence class contains the formula a. Then setting

$$[a] \vee [b]=[a \vee b]$$
$$[a] \wedge [b]=[a \wedge b]$$
$$[a]^1=[a^1]$$

makes $F/\equiv$ into a Boolean algebra. That these operations are well defined, and actually do that is claimed takes some checking and we will not give the details. This Boolean algebra is the classical propositional calculus. If the set V of variables, or atomic formulas, is finite, then $F/\equiv$ is finite, even though F is infinite. It is a fact that if V has n elements. Then $F/\equiv$ has 2^{2^n} elements. If $\{v_1,v_2,\ldots,v_n\}$ is the set of variables, then the elements of the form.

$$w_1\wedge w_2\wedge\ldots\wedge w_n$$

Where is either n or u are called elements, and every element of F is logically equivalent to the join a unique set of monomials. (The element [0] is the join of the empty set of monomials.) Elements written in this fashion are said to be in **disjunctive normal form.**

2.5 A THREE VALUED LOGIC

The construction carried out in the previous section can be generalized in many ways.Perhaps the simplest is to let the set { 0,1} of truth values be larger. Thinking of 0 as representing false and 1 as representing true , we add a third truth value *u* representing undecided. It is common to use ½ instead of u, but a truth value should not be confused *with* a number , so we prefer u. Now proceed as before. Starting with a set of variables, or primitive propositions V build up formulas using this set and some logical connectives. Such logics are called three-valued, for obvious reasons. The set F of formulas is the same as in classical two-valued logic. However, the truth evaluations t will be different, thus leading to a different equivalence relation $\equiv$ on F . There are a multitude of three-valued logics, and their differences arise in the specification of truth tables and implication.

The extending a mapping V → { 0,u,1} to a mapping F → { 0,u,1}, we need to specify how the connectives operate on the truth values. Here is that specification for a particularly famous three-valued logic.

V	0	u	1
0	0	u	1
u	u	u	1
1	1	1	1

∧	0	u	1
0	0	0	0
u	0	u	u
1	0	u	1

	'
)	1
u	u
1	0

Again, we have chosen the basic connectives to be V, ∧, and ',. These operations V and ∧ come simply from viewing { 0,u,1} as the three-element chain with the implied lattice operations. The operation ' is the duality of this lattice. The connectives ⇒ and ⇔ are defined as follows.

⇒	0	u	1
0	1	1	1
u	u	1	1
1	0	u	1

⇔	0	u	1
0	1	u	0
u	u	1	u
1	0	u	1

For this logical system, we still have that a and b are logically equivalent, that is ˜t(a) = ˜ t(b) for all truth valuations t:V → { 0,u,1 } if and only if a ⇔ b is a three-valued tautology.

2.6 Fuzzy Logic

Fuzzy propositional calculus generalizes classical propositional calculus by using the truth set {0,1} . The construction parallel those in the last two sections. The set of building blocks in both cases is a set V of symbols representing atomic or elementary propositions. The set of formulas F is built up from V using the logical connectives ∧ , V. ' (and or , and not , respectively) in the usual way. As in the two-valued and three-valued, propositional calculi, a truth evaluation is gotten by taking any function t: V → [0,1] and extending it to a function t: F → [0,1] by replacing each element a ∈ V which appears in theformula by its value t(a), which is an element in [0,1]. This gives an expression in element of [0,1] and the connectives V, ∧, '. This expression is evaluated by letting

$$x \vee y = \max\{x,y\}$$
$$x \wedge y = \min\{x,y\}$$
$$x' = 1-x$$

for elements x and y in [0,1]. We get an equivalence relation on F by letting two formulas be equivalent if they have the same truth, evaluation for all ˜t. A formula is a tautology if it always has truth value 2. Two formulas u and v . are logically equivalent when ˜t(u) = ˜t(v) for all truth valuations t. As in three valued logic, the law of the excluded middle fails. For an element a ∈ V and a t with t(a) = 0.3, t(a V a') = 0.3 V 0.7 = 0.7 ≠ 1. The set of equivalence classes of logically equivalent formulas forms a kleene algebra, just as in the previous case.

The association of formulas with fuzzy sets in this. With each formula u, associate the fuzzy subset $[0,1]^V$ → [0,1] of [0,1] given by t→ t(u). Thus we have a map from F to $f([0,1]^V)$. This induces a one -to-one mapping from F/≡into the set of mappings from $[0,1]^V$ into [0,1], that is into the set of fuzzy subsets of $[0,1]^V$. This one-to-one mapping associates fuzzy logical equivalence with equality of fuzzy sets.

2.7 Fuzzy and Lukasiewicz logics

The construction of F/≡for the three-valued Lukasiewicz propositional calculus and the construction of the same except for the truth values used. In the first case the set of truth values was {0,u,1} with the tables given, and in the second, the set of truth values was the interval [0,1] with

$$x \vee y = \max\{x,y\}$$
$$x \wedge y = \min(x,y)$$
$$x' = 1-x$$

we remarked that in each case the resulting equivalence classes of formulas formed kleene algebras.

Theorem 1

The propositional calculus for three-valued lukasiewicz logic and the propositional calculus for fuzzy logic are the same

proof. we outline a proof . truth evaluations are mappings f form F into the set of truth values satisfying

$$f(v \vee w) = f(v) \vee f(w)$$
$$f(v \wedge w) = f(v) \wedge f(w)$$
$$f(v') = f(v)'$$

For all formulas v and w in F . Two formulas in F are equivalent if and only if they have the same values for all truth valuations. So we need that two formulas have the same value for all truth valuations into [0,1] if and only if they have the same values for all truth valuations into [0,u,1). First, let $\prod$ be the Cartesian product $\prod_{x\in(0,1)}$ { 0,u,1} with V, ∧ and ' defined componentwise. If two truth valuations from F into $\prod$ differ on an element, then these functions followed by the projection of $\prod$ into one of the copies of { 0.u.1} differ on that element. If two truth valuations from F into {0,u,1} differ on an element, then these two functions followed by is a lattice embedding of [0,u,1} into [0,1] differ on that element. There is a lattice embedding [0,1] → $\prod$ given by y → $\{y_x\}_x$, where y_x is 0,u, or 1 depending on whether y is less than x, equal to x or greater than x. if two truth valuations from F into [0,1] differ on an element, then these two functions followed by this embedding of [0,1] into $\prod$ will differ on that element. The upshot of all this is that taking the truth values to be the lattices [0,u,1}, [0,1], and $\prod$ all induce the same equivalence relation on F, and hence yield the same propositional calculus.

2.8 INTERVAL VALUED FUZZY LOGIC

A fuzzy subset of a set S is a mapping A ; U → [0,1]. The value a(u) for a particular u is typically associated with a degree of belief of some expert. An increasingly prevalent view is that this method of encoding information is inadequate. Assigning an exact number to an expert's opinion is too restrictive. Assigning an interval of values is more realistic. This means replacing the interval [0,1] of fuzzy values by the set [(a,b) ; a,b ∈ [0,1], a≤b}. A standard notation for this set is $[0,1]^{[2]}$. An expert's degree of belief for a particular element u ∈ U will be associated with a pair (a,b) ∈ $[0,1]^{[2]}$. Now we can construct the propositional calculus whose truth values are the elements of $[0,1]^{[2]}$]. But first we need the appropriate algebra of these truth values. It is given by the formulas.

$$(a,b) \vee (c,d) = (a \vee c, b \vee d)$$

$$(a,b) \wedge (c,d) = (a \wedge c, b \wedge d)$$

$$(a,b)^{1} = (b^{1}, a^{1})$$

Where the operations V, ∧ , and ^{1} on elements of [0,1] are the usual ones, commonly referred to in logic to in logic as the disjunction (V), conjunction (Λ) , and negation.

2.9 CANONICAL FORMS :-

As in classical two-valued propositional calculus, every formula that is, every Boolean expression such as $a \wedge (b \vee c) \wedge d^1$ has a canonical form, the well-known disjunctive normal form. For example, the disjunctive normal form for $(a \vee b) \wedge c^1$ in the logic on the variables { a,b,c} is

$$(a \wedge b \wedge c') \vee (a \wedge b' \wedge c') \vee (a' \wedge b \wedge c')$$

and that of $(a \wedge c') \vee (b \wedge c')$ is the same form exactly. Of course, we could have just used the distributive law and noted equality, but that is not the point here. In this classical case, two formulas can be checked for logical equivalence by putting them in their canonical forms and noting whether or not the two forms are identical. Alternately, one can check logical equivalence by checking equality for all truth evaluations of the two expressions. Since the set { 0,1} of truth values is finite, this is a finite procedure.

Now for Lukasiewicz's three-valued logic, which is equal to fuzzy propositional calculus. two formulas may be similarly tested for logical equivalence, that is, by checking equality of all truth evaluations. Two formulas in fuzzy propositional calculus are logically equivalent if any only if they are logically equivalent in Lukasiewicz's three-valued propositional calculus.

The normal form for De Morgan algebras stems from realizing that all conjunctions of literals as well as I, are join irreducible. The normal form for Boolean algebras stems from realizing that the only join irreducible elements in the Boolean case are the complete conjunction of literals in which each variable occurs exactly once. For example, if the variables are x_1,x_2,x_3 then $x_1 \wedge x_2 \wedge x_3 \wedge$ and $x_1 \wedge x'_2 \wedge x'_3$ are complete disjunctions while $x_1 \wedge x'_2$ and $x_2 \wedge x'_3$ are not. The empty disjunction is 0 and the disjunction of all the complete conjunctions is 1.

The join irreducibles in the kleene case are a little more subtle. R the variables are x_1, x_2, x_3, ...x_n, then a conjunction of literals is join irreducible if any only if it is 1 , or it contains at most one of the literals for each variable, or it contains at least one of the literals for each variable, or it contains at least one of the literals for each variable. Suppose n = 3. Here are some examples.

1. $x_1 \wedge x_2 \wedge x_3$ is join irreducible. It contains at least one of the literals for each variable. (It also contains at most one of the literals for each variable, so qualifies on two counts).
2. $x_1 \wedge x_2 \wedge x'_3$ is join irreducible for the same reasons as above.
3. $x_1 \wedge x_2 \wedge x'_3$ is join irreducible. It does not contain at least one of the literals for each variable, and it contains two literals for the variable x_2.
4. $x_1 \wedge x'_1 \wedge x_2 \wedge x'_2$ is join irreducible. It contains at least one of the literals for each variable.
5. $x_1 \wedge x'_1 \wedge x_2 \wedge x'_2$ is not join irreducible. It does not contain at least one of the literals for each variable, and it contains two literals for two variables.
6. $x_1 \wedge x_2$ is join irreducible. It contains at most one of the literals for each variable.
7. x_3 is join irreducible. It contains at most one of the literals for each variable.

Now the normal form for the Boolean algebra case, that is, for F_1,is of course well-known: every element is uniquely an disjunction of complete conjunctions of literals. Instead of getting into this, we will describe the procedure for putting an arbitrary formula in Kleene normal form. In the examples illustrating the steps, we assume that are three variables.x_1,x_2,x_3.

1. Given an formula w, first use De Morgan's laws to move all the negation in, so that the formula is rewritten as an formula w1 which is just meets and joints of the literals, 0, and 1. For example, x_1^(x'_2^x_3)'would be replaced by X_1^(X_2vX'_3).
2. Next use the distributive law to obtain an new formula w_2 from w_1 which is an disjunction of conjunctions involving the literals, 0, and 1. For example, replace x_1^(x_2vx'_3) by (x_1^x'_3). At this point , discard any conjunction in which 0 or a' appears as one of the conjunction, as well as 1 and 0' from any conjunction in which they do not appear alone(if an conjunction consists entirely of 1's and 0',s, then replace the whole thing by 1) This yield an formula w_3
3). Now discard all no-maximal conjuctions among the conjuctions that w_3 is an disjunction of. The type of conjuctions we now are dealing with are either conjuctions of literals or 1 by itself. Of course 1 is above all the others and one conjuction of literals is belowanoher if and only if the former contains all the literals contained in the latter. This process yields an formula w_4.
4. At this point, replace any conjuction of literals, calculate, which contains both literals for at least one variable by the disjunction of all the conjuction of literals for each variable not occurring in c. For example, if one of the conjuctions ids x_1^x'_1^1^X_2, replace it by the disjunction(X_1^x'1^x_2^x_3). (x_3 is the only variable not occuring in x_1^x'1^x_2)
5. Finally, again discard all non-maximal conjuctions among the conjuctions that are left, and if no conjuctions are left, then replace the formula by 0. The formula thus obtained is now in the normal form described above.

We illustrate the Kleen normal form with the two equivalent expressions.

W = A^ (A'ΛB)v(A'ΛB')v(A'ΛC)

W'=AΛA'

In the variables, A,B and C

1. There is nothing to do in this step
2. Applications of the distributive law lead to disjuctions of conjuctions involving the literals.

W_2 = (AΛA^1ΛB)V(AΛ$A^{1\wedge}B^1$)V($A^\wedge A^{1\wedge}C$)

$W'_2 = A^\wedge A^1$

3. Neither of the expressions in # 2 contains any non maximal conjunctions, so w_3=w_2 and $w^1{}_3=w^1{}_2$.
4. Replace

$A^\wedge A^{1\wedge}B$ by ($A^\wedge$A'b^C) v ($A^\wedge A^{1\wedge}B^\wedge C^1$)

$A^\wedge$A'^B' by ($A^\wedge$A'^$B^{1\wedge}$C) v ($A^\wedge A^{1\wedge}B^{1\wedge}c^1$)

$A^\wedge A^{1\wedge}$C by ($A^\wedge A^{1\wedge}$c^B) V (A^A1^C^B')

And

$A^\wedge A^1$ by ($A^\wedge A^{1\wedge}$B^C)V($A^\wedge A^{1\wedge}B^{1\wedge}$C)

V($A^\wedge A^1$B^C') V ($A^\wedge A^{1\wedge}$B^^c^1)

To get

$W_4 = (A \wedge A^1 \wedge B \wedge C) V (A \wedge A^1 \wedge B \wedge C')$
$V(A \wedge A' \wedge B' \wedge C) \; V \; (A \wedge A^1 \wedge B^1 \wedge C^1)$
$V(A \wedge A^1 \wedge C \wedge B) \; V \; (A \wedge A' \wedge C \wedge B^1)$

$W_4 = (A \wedge A^1 \wedge C) \; V \; (A \wedge A^1 \wedge B^1 \wedge C)$
$V(A \wedge A^1 \wedge B \wedge C^1) \; V \; (A \wedge A' \wedge B^1 \wedge C^1)$

5. Discarding all non-maximal conjunctions amoung the conjunctions that are left means in this case, simply discarding repetitions, leading to the normal forms.

$W_5 = (A \wedge \text{-}AB \wedge C) \; V \; (A \wedge \text{-}A \wedge B \wedge C^{11}) \; V \; (A \wedge \text{-}A \wedge B^1) \; V \; (A \wedge B^1 \wedge C^1)$

$w_5 = (A \wedge \text{-}ABC) V \quad (A\text{-}AB^1C) \; V \; (A\text{-}ABC^1) \; V \; (A\text{-}AB^1C)$

EXCERCISES

1. Show that there are fuzzy quantities A and B, such that
   (a) $A-A \neq 0$
   (b) $(A+B)-B \neq A$
   (c) $A/A \neq 1$
   (d) $A/B \; B \neq A$
2. Show that for fuzzy quantities, multiplication does not distribute over addition. That is, $A(B+C) \neq A B + AC$.
3. Let S and T be closed and bounded subsets of R. Show that $(X_s/X_r)(u) = X_s(u \; x) \wedge X_r(x)$ for some x.
4. Compute the α-cuts of the sum of two triangular numbers.
5. For f :R→R and A∈f(R), write down the membership function of f(A) when
   $f(x)=-x, \qquad f(x)=x^2$
   $f(x)=x^5, \qquad f(x)=|x|$
6. Define the fuzzy quantities A and B by
   $A(x)=1/2(1+e^{-x/2})$
   $B(x)=1$

   Show that A and B are convex, A +B is convex, but $(A+B)_{3/4} \neq A_{3/4}+B_{3/4}$.
7. Write down the tables for => and for classical two-valued propositional logic.
8. In two-valued propositional calculates, verify that two propositions a and a and b are logically equivalent if and only a => b is a tautology.
9. We write a = b for a <=>. Verify the following for two-valued propositional calculates.
   (a) $a'' = a$
   (b) $a \; V \; a^1 = 1$
   (c) $a \wedge a^1 = 0$
   (d) $a = a \; V \; a$
   (e) $a \; V \; b = b \; V \; a$

(f) $a \wedge b = b \wedge a$

(g) $a V (bV c) = (a V b) V c$

(h) $a\wedge (b\wedge c) = (a \wedge b) \wedge c$

(i) $a\wedge (b V c) =(a \wedge b) V(a\wedge c)$

(j) $a V(b\wedge c) = a V b) V (a \wedge c)$

(k) $(a V b)^1 = a' \wedge b'$

(l) $(a\wedge b)^1 = a^1 v b^1$

10. In Bochvar'soni three valued logic, <=> is defined by

<=>	0	u	1
0	1	u	0
u	u	u	u
1	0	u	1

Verify that a and b being logically equivalent does not imply that a <=> b is a three valued tautology.

11. Show that u V u = u is changed to uVu=1 in the table for V in Lukasiewicz'soni three-valued logic, then the law of the exclued middle holds.

12. Let a be a formula in fuzzy logic. Show that if $t(aVa^1) = 1$, then necessarily $t(a) \in \{0,1\}$.

13. Show that {0,u,1} with 0<u<1 is a Kleene algebra. For any set S, Show that {0v,1} S is a Kleene algebra.

14. Show that in the algebra ([0,1], V, Λ,^{1},0,1) the inequality $X^\wedge X^1 \le \gamma V \gamma' 1$ holds for all x and γ in [0,1]. Show that this inequality does not hold in ([0,1][2].V,,',0,1)

15. Show that

$$A^\wedge ((A^{1\wedge}B)V(A^{1\wedge}B^1)V(A^{1\wedge}C)) = A\ ^\wedge A^1$$

Is false for fuzzy sets taking values in [0,1][2]

16. In the three variables A,B,C find the disjuctive normal firm, the Kleene normal form, and the De Morgan normal form for

(a) $AV (A^1 \Lambda\ b \Lambda B^1)$

(b) $A\Lambda(B\ VC)^1$

UNIT - III

DISTRIBUTIONS OF RANDOM VARIABLES

TABLE OF CONTENTS

INTRODUCTION :

Many kinds of investigations may be characterised in part by the fact that repeated experimentation, under essentially the same conditions, is more or less standard procedure . Each experiment terminates with an *outcome.* But it is characteristic of these experiments that the outcome cannot be predicted with certainty prior to the performance of the experiment.

Suppose that we have such an experiment, the outcome of which cannot be predicted with certainty, but the experiment is of such a nature that the collection of every possible outcome can be described prior to its performance. If this kind of experiment can be repeated under the same conditions, it is called ***a random experiment***, and the collection of every possible outcome is called the experimental space or the ***sample space.***

Example1. In the toss of a coin, let the outcome tails be denoted by T and let the outcome heads be denoted by *H.* If we assume that the coin may be repeatedly tossed under the same conditions, then the toss of this coin is an example of a random experiment in which the outcome is one of the two symbols T and H; that is the sample space is the collection of these two symbols.

3.1 ALGEBRA OF SETS

Definition 1. If each element of a set A_1 is also an element of set A_2, the set A_1 is called a subset of the set A_2. This is indicated by writing $A_1 \subset A_2$. If $A_1 \subset A_2$ and also $A_2 \subset A_1$, the two sets have the same elements, and this is indicated by writing $A_1=A_2$.

Definition 2. If a set A has no elements, A is called the null set. This is indicated by writing $A=\emptyset$.

Definition 3. The set of all elements that belong to at least one of the sets A_1 and A_2 is called the union of A_1 and A_2. The union of A_1 and A_2 is indicated by writing $A_1 \cup A_2$.

Definition 4. The set of all elements that belong to each of the sets A_1 and A_2 is called the intersection of A_1 and A_2. The intersection of A_1 and A_2 is indicated by writing $A_1 \cap A_2$.

Definition 5. In certain discussions or considerations the totality of all elements that pertain to the discussion can be described. This set of all elements under consideration is given a special name. It is called the space. We shall often denote spaces by capital script such as A, B, and C.

Definition 6. Let A denote a space and let A be a subset of the set A . The set that consists of all elements of A that are not elements of A is called the complement of A . The complement of A is denoted by A* (In particular. A * = ∅).

Example. Given $A \subset A$. Then $A \cup A^* = A$ $A \cap A^* = \varnothing$, $A \cup A = A$, $A \cap A = A$, and $(A^*)^* = A$.

3.2 SET FUNCTIONS:

In the calculus, functions such as

$f(x) = 2x, \ -\infty < x < \infty,$

or

$g(x,y) = e^{-x-y}, \ 0 \ x < \infty, 0 < y < \infty$, or possibly

$h(x_1, x_2, \ldots x_n) = 3x_1x_2 \ldots x_n, \ 0 \le x_i \le 1, i = 1,2,\ldots,n,$

$= 0$ elsewhere,

were of common occurrence. The value of f (x) at the "Point x =1 " is f(1) =2; the value of g(x,y) at the "Point (-1,3)" is f(-1,3) = 0;

the value of $h(x_1,x_2 \ldots x_n)$ at the "Point (1,1,.....1)" is 3. Functions such as these are called functions of a point or, more simply, Point functions.

Notations :

The symbols $\int_A f(x)dx$

will mean the ordinary (Riemann) integral of f(x) over a prescribed one-dimensional set A; the symbol

$\int_A \int g(x,y)\, dxdy$

will mean the Riemann integral of g(x,y) over a prescribed two-dimensional set A; and so on.

Example. Let A be a one-dimensional set and let

$Q(A) = \int_A e^{-x} dx$

Thus, if $A = \{x; 0 \le x < \infty\}$, then

$Q(A) = \int_0^\infty e^{-x} dx = 1;$

if $A = \{x; 1 \le x \le 2\}$, then

$Q(A) = \int_1^2 e^{-x} dx = e^{-1} - e^{-2};$

if $A_1 = \{x; 0 \le x \le 1\}$ and $A_2 = \{x; 1 < x \le 3\}$, then

$Q(A_1 \cup A_2) = {}_0\int^{3} e^{-x}\, dx$

$= {}_0\int^{1} e^{-x}\, dx + \int 31\ e\text{-}x\ dx$

$= Q(A_1) + Q(A_2);$

if $A = A_1 \cup A_2$, where $A_1 = \{ x; 0 < x < 2 \}$ and $A_2 = \{ x; 1 < x < 3\}$, then

$Q(A) = Q(A_1 \cup A_2) = {}_0\int^{3} e^{-x}\, dx$

$= {}_0\int^{2} e\text{-}x\, dx + {}_1\int^{3} e^{-x}\, dx - {}_1\int^{2} e^{-x}\, dx$

$= Q(A_1) + Q(A_2) - Q(A_1 \cap A_2).$

Example . Let A be a set in n dimensional space and let

$Q(A) = \int \ldots_A \int dx_1\, dx_2 \ldots\ldots dx_n$

If $A = \{ (x_1, x_2 \ldots\ldots x_n);\ 0 < X_1 < x_2 < \ldots < x_n < 1 \}$, then

$Q(A) = {}_0\int^{1} {}_0\int^{x_n} \ldots {}_0\int^{x_3} dx_1\, dx_2 \ldots dx_{n-1}\, dx_n$

$= 1/n!$, where $n! = n(n-1) \ldots 3.2.1.$

3.3 THE PROBABILITY SET FUNCTION.

Let $\mathcal{C}$ denote the set of every possible outcome of a random experiment; define a set function P(C) such that if C is a subset of $\mathcal{C}$, then P(C) is the probability that the outcome of the random experiment is an element of C.

Definition : If P(C) is defined for a type of subset of the space $\mathcal{C}$, and if

(a) $P(C) \ge 0$,

(b) $P(C_1 \cup C_2 \cup C_3 \cup \ldots) = P(C_1) + P(C_2) + P(C_3) + \ldots$ where the sets C_i, $i = 1,2,3, \ldots$ are such that no two have a point in common, (that is, where $C_i \cap C_j = \varnothing$, $i = j$).

(c) $P(C) = 1$,

then P(C) is called the probability set function of the outcome of the randodm experiment.

Theorem 1.

For each $C \subset \mathcal{C}$, $P(C) = 1 - P(C^*)$.

Proof. We have $\mathcal{C} = C \cup C^*$ and $C \cap C^* = \varnothing$ By definition, it follows that $1 = P(C) + P(C^*)$, Hence , $P(c) = 1 - p(c^*)$.

Theorem 2:

The probability of the null set is zero; that is $P(\varnothing) = 0$.

Proof. In Theorem 1, take $C = \varnothing$ so that $C^* = \mathcal{C}$. Accordingly, we have

$P(\varnothing) = 1 - P(\mathcal{C}) = 1 - 1 = 0,$

Theorem 3.

If C_1 and C_2 are subsets of C such that $C_1 \subset C_2$, then $P(C_1) < P(C_2)$.

Proof. Now $C_2 = C_1 \cup (C^*_1 \cap C_2)$ and $C_1 \cap (C^*_1 \cap C_2) = \varnothing$. Hence, from (b) of definition,

$P(C_2) = P(C_1) + P(C^*_1 \cap C_2)$

However, from (a) of Denifition $P(C^*_1 \cap C_2) \geq 0$; accordingly, $P(C_2) \geq P(C_1)$

Theorem 4.

For each $C \subset \mathbf{C}$, $0 \leq P(C) \leq 1$

Proof. Since $\varnothing \subset C \subset \mathbf{C}$, we have by Theorem 3 that

$P(\varnothing) < P(C) \leq P(\mathbf{C})$ or $0 \leq P(C) \leq 1$ the desired result,

Theorem 5.

If C_1 and C_2 are subsets of $\mathbf{C}$ then $P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2)$

Proof . Each of the sets $C_1 \cup C_2$ and C_2 can be represented, respectively, as a union of nonintersecting sets as follows

$C_1 \cup C_2 = C_1 \cup (C^*_1 \cap C_2)$ and $C_2 = (C_1 \cup C_2) \cup (C^*_1 \cap C_2)$

Thus, from (b) of Definition

$$P(C_1 \cup C_2) = P(C_1) + P(C^*_1 \cap C_2)$$

And

$$P(C_2) = P(C_1 \cap C_2) + P(C^*_1 \cap C_2).$$

If the second of these equations is solved for $P(C^*_1 \cap C_2)$ and this result substituted in the first equation, we obtain.

$$P(C_1 \cup C_2) = P(C_1) + P(C_2) - P(C_1 \cap C_2).$$

Example : Two coins are to be tossed and the outcome is the ordered pair (face on the first coin, face on the second coin). Thus the sample space may be represented as $\mathbf{C}$ = (c:c = (H,H), (H,T), (T,H), (T,T)}. Let the probability set function assign a probability of ¼ to each element of $\mathbf{C}$ Let C_1 = {c : c H,H), (H,T)} and C_2 = {c;c = (H,H), (T,H). Then $P(C_1) = P(C_2) = ½$ $P(C_1 \cap C_2) = ¼$ and in accordance with Theorem 5, $P(C_1 \cup C_2) = ½+½-¼=¾$.

3. 4 RANDOM VARIABLES (r.v)

Let the random experiment be the toss of a coin and let the sample space associated with the experiment be $\mathbf{C}$ = { c; where c is T or c is H} and T and H represent, respectively, tails and heads. Let X be a function such that X(c) = 0 if c is T and let X(c) = 1 if c is H. Thus X is a real-valued function defined on the sample space C which takes us from the sample space C to a space of real number **A** = { x;x = 0,1}.

Definition

Given a random experiment with a sample space $\mathbf{C}$. A function X, which assigns to each element $c \in \mathbf{C}$ one and only one real number X (c) = x, is called a random variable. The space of X is the set of real numbers **A** = $\{x; x = X(c), c \in \mathbf{C}\}$.

Definition

Given a random experiment with a sample space $\mathbf{C}$. Consider two random variables X_1 and X_2 which assign to each element c of $\mathbf{C}$ one and only one ordered pair of numbers $X_1(c) = x_1, X_2(c) = x_2$. The space of X_1 and X_2 is the set of ordered pairs **A** = $\{ (x_1,x_2);\ x_1=X_1(C), x_2 = X_2(C), C \in \mathbf{C} \}$.

Definition

Given a random experiment with the sample space **C**. Let the random variable X_i assign to each element $c \in$ **C** one and only one real number $X_i(c) = x_i$, $i = 1,2,\ldots,n$. The space of these random variables is the set of ordered n-tuplets **A** $= \{(x_1, x_2 \ldots x_n);\ x_1 = X_1(c) \ldots,\ x_n = X_n(c),\ c \in$ **C** $\}$. Further, let A be a subset of **A**. Then $\Pr[(X_1, \ldots, X_n) \in A] = P(C)$, Where $C = \{c; c \in$ **C** and $[X_1(c), X_2(c), \ldots X_n(c)] \in A\}$

Example of a sample space **C** an interval.

Example.

Let the outcome of a random experiment be a point on the interval (0,1). Thus,

C $= \{c; 0 < c < 1\}$. Let the probability set function be given by

$P(C) = \int_C dz$

For instance, if $C = \{c; ¼ < c < ½\}$, then

$P(C) = \int_{1/4}^{1/2} dz = ¼$.

Define the random variable X to be $X = X(c) = 3c + 2$. Accordingly, the space of X is **A** $= \{x; 2 < x < 5\}$. We wish to determine the probability set function of X, namely $P(A)$, $A \subseteq$ **A**. At this time, let A be the set $(x; 2 < x < b)$, where $2 < b < 5$. Now $X(c)$ is between 2 and b when and only when $c \in C = \{c; 0 < c < (b-2)/3\}$. Hence

$$P_x(A) = P(A) = P(C) = \int_0^{(b-2)/3} dz.$$

In the integral, make the change of variable $x = 3z + 2$ and obtain

$$P_x(A) = P(A) = \int_2^b 1/3\,dx = \int_A 1/3\,dx.$$

Since $A = \{x; 2 < x < b\}$. This kind of argument holds for every set $A \subset$ **A** for which the integral

$$P(A) = \int_A 1/3\,dx$$

exists. Thus the probability set function of X is this integral.

Example

Let the probability set function P(A) of a random variable X be

$P(A) = \int_A f(x)\,dx$, where $f(x) = 3x^2/8$, $x \in$ **A** $= \{x; 0 < x < 2\}$.

Let $A_1 = \{x; 0 < x < 1/2\}$ and $A_2 = \{x; 1 < x < 2\}$ be two subsets of **A**. Then

$$P(A_1) = pr(X \in A_1) = \int_{A_1} f(x)dx = \int_0^{1/2} 3x^2/8\,dx = 1/64$$

and

$$P(A_2) = pr(X \in A_2) = \int_{A_2} f(x)dx = \int_1^2 3x^2/8\,dx = 7/8.$$

To compute $P(A_1 \cup A_2)$, we note that $A_1 \cap A_2 = \varnothing$; then we have $P(A_1 \cup A_2) = P(A_1) + P(A_2) = 57/64$.

Example

Let A $=\{(x,y);0<x<y<1\}$ be the space of two random variables X and Y. Let the probability set function be

$$P(A)=\int_A\int 2\,dx\,dy.$$

If A is taken to be $A_1=\{(x,y);1/2<x<y<1\}$, then

$$P(A_1) = Pr[(X,Y) \in A_1] = \int_{1/2}^{1}\int_{1/2}^{y} 2\,dx\,dy = 1/4$$

If A is taken to be $A_2=\{(x,y);\ x<y<1, 0<x \le 1/2\}$, then $A_2=A^*_1$, and

$$P(A_2) = Pr[X,Y] \in A_2] = P(A^*_1) = 1-P(A_1) = 3/4$$

3.5 THE PROBABILITY DENISTY FUNCTION

Let X denote a random variable with space **A** and let A be a subset of **A** . If we know how to compute P(C), C ⊂ C, then for each A under consideration we can compute P(A) = Pr (X ∈ A); that is, we know how the probability is distributed over the various subsets of **A** .

In this section, we discuss some random variables whose distributions can be described very simply by what will be called the probability density function.

(a) THE DISCRETE TYPE OF RANDOM VARIABLE:

Let X denote a random variable with one-dimensional space **A** . Suppose that the space **A** is a set of points such that there is at most a finite number of points of **A** in every finite interval. Such a set **A** will be called a set of discrete points. Let a function f(x) be such that $f(x) > 0, x \in$ **A** , and that

$$\sum_{A} f(x) = 1.$$

Whenever a probability set function P(A) , A ⊂ **A** , can be expressed in terms of such an f(x) by

$$P(A) = Pr(X \in A) = \sum_{A} f(x),$$

Then X is called a random variable of the discrete type, and X is said to have a distribution of the discrete type.

Example

Let X be a random variable of the discrete type with space A = {x;x = 0,1,2,3,4} Let

$$P(A) = \sum_{A} f(x),$$

Where

$$F(x) = \frac{4!}{X!\,(4-x)!}\,(1/2)^4,\quad x \in A$$

And as usual, 0! = 1 Then if A = { x:x = 0, 1 } we have

$$Pr(X \in A) \frac{4!}{X!(4-x)!} (\tfrac{1}{2})^4 + \frac{4!}{1!3!} (\tfrac{1}{2})^4 = \frac{5}{16}$$

(b) THE CONTINUOUS TYPE OF RANDOM VARIABLE:

Let the one dimensional set **A** be such that the Riemann integral

$$\int_A f(x)\,dx = 1,$$

where (1) f(x) > 0, x ∈ **A**, and (2) f(x) has at most a finite number of discontinuities in every finite interval that is a subset of **A** . if **A** is the space of the random variable X and if the probability set function p(A), A ⊂ A, can be expressed in terms of such an f(x) by

$$P(A) = pr(X \in A) = \int_A f(x)\,dx,$$

Then X is said to be a random variable of the continuous type and to have a distribution of that type.

Example : Let the Space A = { x;0 < x < ∞}, and let

$$f(x) = e^{-x}, \quad x \in A.$$

If X is a random variable of the continuous type so that

$$Pr(X \in A) = \int_A e^{-x}\,dx,$$

We have, with A = {x;0 < x < 1},

$$Pr(X \in A) = \int_0^1 e^{-x}\,dx = 1 - e^{-1}$$

Note that pr (X ∈ **A**) is the area under the graph of $f(x) = e^{-x}$ which lies above the x-axis and between the vertical lines x =0 and x=1.

If two probability density functions of random variables of the continuous type differ only on a set having probability zero, the two corresponding probability set functions are exactly the same. Unlike the continuous type, the P.d.f. of a discrete type of random variable may not be changed at any point since a change in such a p.d.f. alters the distribution of probability. If a p.d.f in one or more variables is explicitly defined, we can see by inspection whether the random variables are of the continuous or discrete type. For example, it seems obvious that the p.d.f.

$$F(x,y) = \frac{9}{4x+y}, \quad x = 1,2,3,\ldots,y = 1,2,3,\ldots,$$

= 0 elsewhere

is clearly a p.d.f. of two continuous-type random variables X and Y.

Example : Let the random variable X have the p.d.f.

$f(x)=2x, 0<x<1,$

$=0$ elsewhere.

Find Pr(1/2<X<3/4) and Pr(-1/2<X1/2). First,

$Pr(1/2<X<3/4)={}_{3/4}\int^{½} f(x)dx = {}_{3/4}\int^{½} 2xdx=5/16.$

Now,

$$Pr(-1/2<X<1/2) = \int_{-½}^{½} f(x)\,dx$$

$$= \int_{-½}^{0} d_x + \int_{0}^{½} 2x\, d_x$$

$$= 0+1/4$$

$$= ¼$$

Example : Let $f(x,y) = 6x^2y, 0<x<1, 0<y<1,$

= 0 elsewhere,

be the p.d.f. of two random variables X and Y. We have, for instance, pr(0<X<3/4,1/3<Y<2) = ${}_{1/3}\int^{2} {}_{0}\int^{3/4} f(x,y)\,dx\,dy$

$$= {}_{1}\int^{1/3} {}_{0}\int^{3/4} 6x^2y\,dx\,dy + {}_{1}\int^{2} {}_{0}\int^{3/4} dx\,dy$$

$=3/8+0=3/8.$

Now that this probability is the volume under the surface $f(x,y)=6x^2y$ and above the rectangular set{(x,y);0<x<3/4,1/3<y<1} in the xy-plane.

3.6 THE DISTRIBUTION FUNTION

Let the randome variable X have the probability set function P(A), where A is a one-dimensional set. For all such sets A we have $P(A)=Pr(X\in A)=Pr(X\le x)$. This probability depends on the point x; This point function is denoted by the symbol $F(x)=Pr(X\le x)$. The function F(x) is called the distribution function (sometimes, cummulative distribution function) of the random variable X. Since

$F(x) = Pr(X\le x)$, then, with f(x) the p.d.f., we have

$F(x) = \sum_{w \le x} f(w),$

For the discrete type of random variable, and

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

For the continuous type of random variable.

Example 1:

Let the random variable X of the discrete type have the p.d.f. f(x)=x/6,x = 1,2,3,zero elsewhere. The distribution function of X is

$F(x) = 0, \quad x<1,$

$= 1/6, \quad 1 \le x<2,$

$= 3/6, \quad 2 \le x<3,$

$= 1, \quad 3 \le x.$

Here, F(x) is a step function that is constant in every interval not cantaining 1, 2,, or 3, but has steps of heights, 1/6, 2/6 and 3/6 at those respective points. It is also seen that F(x) is everywhere continuous to the right.

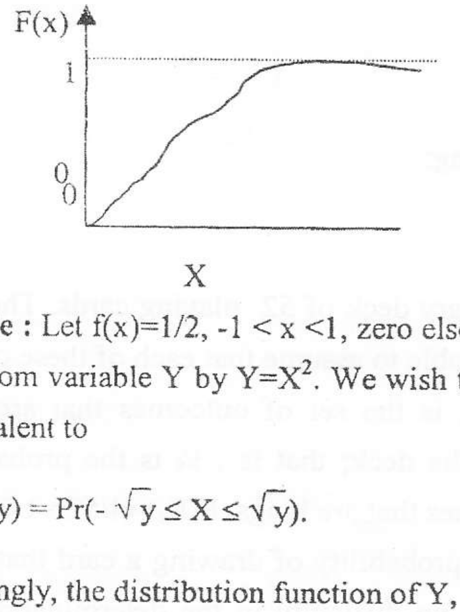
Example :

Let the random variable X of the continous type have the p.d.f. f(x)=2/x3,1 < x<∞, zero elsewhere. The distribution function of X is

$F(x) = \int_{-\infty}^{x} 0\, dw = 0, \; x<1,$

$= \int_{1}^{x} 2/w^3\, dw = 1 - 1/x^2, \; 1 \le x.$

The graph of this distribution function is depicted in Figure



Example : Let f(x)=1/2, -1 < x <1, zero elsewhere, be the p.d.f. of the random variable X. Define the random variable Y by $Y=X^2$. We wish to find the p.d.f. of Y. If y≥0, the probability Pr(Y≤y) is equivalent to

$Pr(X^2 \le y) = Pr(-\sqrt{y} \le X \le \sqrt{y}).$

Accordingly, the distribution function of Y, G(y)= Pr(Y≤y), is given by

$G(y)=0, y<0,$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} \tfrac{1}{2}\, d_x = \sqrt{y},\ 0\le y<1,$$

$=1\ ,\ 1\le y.$

Since Y is a random variable of the continuous type, the p.d.f. of Y is g(y)=G'(y) at all points of continuity of g(y). Thus we may write

$$G(y) = \frac{1}{2\sqrt{y}}\ ,\quad 0<y<1,$$

0 elsewhere.

Let the random variables X and Y have the probability set function P(A), where A is a two -dimensional ser. If A is the unbounded set{(u,v); u≤x, v≤y}, where x and y are real numbers, we have

$P(A)=Pr[(X,Y)\in A]= Pr(X\le x,Y\le y).$

This function of the point(x, y) is called the distribution function of X and Y and is denoted by

$F(x,y) = Pr(X \le x, Y \le y).$

If X and Y are random variables of the continuous type that have p.d.f. f(x,y), then

$$\int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)du\, dv.$$

3.7 PROBABILITY MODELS

The probability model described in the following:

Example

Let a card be drawn at random from a ordinary deck of 52 playing cards. The sample space is the union of k=52 outcomes, and it is reasonable to assume that each of these outcomes has the same probability 1/52 . Accordingly, if E_1, is the set of outcomes that are spades, P(E1)=13/52=1/4 because there are r_1=13spades in the deck; that is , ¼ is the probability of drawing card that is a spade. If E_2 is the set of outcomes that are kings, $P(E_2)$=4/52= 1/3 because there are r_2=4 kings in the deck; that is , 1/13 is the probability of drawing a card that is king. These computations are very easy because there are no difficulty in the determination of the appropriate values of r and k. However, instead of drawing only one card, suppose that five card are taken, at random and without replacement, from this deck. We can think each five card hand as being outcome in a sample space. It is responsible to assume that each of these outcomes as the same probability. Now if E_1 is the set of outcomes in each card of the hand is a spade $P(E_1)$ is equal to the number r1 of all spade hands divided by the total number, say k, five-card hands. It is shown in many books on algebra that

$$r_1 = {}^{13}C_5 = \frac{13!}{5!\,8!} \text{ and } k = {}^{52}C_5 = \frac{52!}{5!\,47!}.$$

In general, if n is a positive integer and if x is a non negative integer with nC_x then the binomial coefficient

$$nc_x = n!/x!(n-x)!$$

is equal to the number of combinations of n things taken x at time. Thus, here,

$$P(E_1) = \frac{13C_5}{52C5} = \frac{13.12.11.10.9}{52.51.50.49.48} = 0.0005$$

approximately. Now, let E_2 be the set of outcomes in which atleast one card is a spade. Then E_2^* is the set of outcomes in which no card is a spade.

There are $r_2^* = {}^{39}C_5$ such out comes Hence

$P(E_2^*) = \frac{{}^{39}C_5}{{}^{52}C_5}$ and $P(E_2) = 1 - P(E_2^*)$.

Now suppose that E_3 is the set of outcomes in which exactly three cards of kings and exactly two card are queens.

We can set the three kings in any one of the $4C_3$ ways and the queens in any one of ${}^4C_2$ ways by a well-known counting principle, the number of outcomes in E_3 is $r_3 = {}^4C_3\,{}^4C_2$ Thus $P(E_3) = {}^4C_3\,{}^4C_2/{}^{52}C_5$. Finally, let E_4 be the set of outcomes in which there are exactly two kings, two queens, and one jack. Then

$$P(E_4) = {}^4C_2\,{}^4C_2\,{}^4C_1/{}^{52}C_5$$

because the numerator of this fraction is the number of outcomes in E_4.

3.8 MATHEMATICAL EXPECTATION

Let X be a random variable having a p.d.f. f(x) and let u(X) be a function of X such that

$E[u(x)] = \int_{-\infty}^{\infty} u(x) f(x) dx$, exist is

if X is a continuous type of random variable,

And, $E[u(x)] = \sum_x u(x) f(x)$

exists, if X is a discrete type of random variable. The integral, or the sum, as case may be, is called the mathematical expectation.

Remarks.

The usual definition of E[u(X)] requires that the integral(or sum) converge absolutely.

We may observe that u(X) is a random variable Y with its own distribution of probability. Suppose the p.d.f. of Y is g (y). Then E (Y) is given by

$\int_{-\infty}^{\infty} yg(y)\,dy$ or $\sum_y yg(y)$, according as Y is of the continuous type or of the discrete type.

Results:

(a) If k is a constant, then E (k) = k.

(b) If k is a constant and v is a function, then E(kV) = kE(v).

(c) If k_1 and k_2 are constants and v_1 and v_2 are functions, then $E(k_1v_1+k_2v_2) = k_1E(v_1)+k_2E(v_2)$.

Example1:

Let X have the p.d.f.

F (x) = 2(1-x), 0<x<1,

= 0 elsewhere.

Then

$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_0^1 (x)2(1-x)\,dx = 1/3,$

$E(x^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_0^1 (x^2)2(1-x)\,dx = 1/6,$

And, of course,

$E(6X+3X^2) = 6(1/3)+3(1/6) = 5/2.$

Example 2:

Let X have the p.d.f.

f(x)=x/6, x=1,2,3,

=0 elsewhere.

Then

$E(X^3) = \sum_x x^3 f(x) = \sum_{x=1}^{3} x^3 \frac{\Pi}{6}$

$= 1/6+16/6+81/6=98/6.$

Example 3:

Let X and Y have a p.d.f.

$F(x, y) = x+y, 0<x<1, 0<y<1,$

$= 0$ elsewhere.

Then,

$E(XY^2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy^2 f(x,y)\, dx\, dy$

$= \int_0^1 \int_0^1 xy^2(x+y)\, dx\, dy$

$=17/72.$

Example 4:

Let us divide, at random, a horizandal line segment of length 5 into two parts. If X is the length of the left-hand part, it is reasonable to assume that X has the p.d.f.

$F(x) = 1/5, \quad 0<x<5,$

$= 0$ elsewhere.

The expected value of the length X is E(X)= 5/2 and the expected value of the length 5-X is E(5-X)=5/2. But the expected value of the product of the two length is equal to

$E[X(5-X)] = \int_0^5 x(5-x)(1/5)\, dx = 25/6 \neq (5/2)^2.$

That is, in general, the expected value of the product is not equal to the product of the expected values.

3.9 Some Special Mathematical Expectations:

Let $u(X)=X$, where X is a random variable of the discrete type having a p.d.f. $f(x)$. Then

$$E(X)=\Sigma_x xf(x).$$

If the discrete points of the space of positive probability density are $a_1,a_2,a_3,\ldots$,then

$$E(X) = a_1 f(a_1)+a_2 f(a_2)+a_3 f(a_3)+\ldots$$

This sum of products is seen to be a "weighted average" of the values $a_1,a_2,a_3,\ldots$, the "weight" associated with each ai being f(ai). This suggest that we call E(X) the arithmetic mean of the values of X, or , more simply, the mean value of X(or the mean value of the distribution).

The mean value μ of a random variable X is defined, when it exists, to be $\mu = E(X)$, where X is a random variable of the discrete or of the continuous type .

The variance of X will be denoted by σ^2, and we define $\sigma^2 = E[(X-\mu)^2]$,whether X is a discrete or a continuous type of random variable.

It is worthwhile to observe that

$$\sigma^2 = E[(X-\mu)^2]=E(X^2-2\mu X+\mu^2);$$

and since E is a linear operator,

$$\sigma^2 = E(X^2)-2\mu E(X)+\mu^2$$

$$= E(X^2)-2\mu^2+\mu^2$$

Result : $\sigma^2 = E(X^2)-\mu^2$.

Example 1. Let X have the p.d.f.

$$F(x) = \frac{1}{2(x+1)}, \quad -1<x<1,$$

$$= 0 \text{ elsewhere}$$

Then the mean value of X is

$\mu = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{-1}^{1} x \frac{x+1}{2} dx = 1/3$

while the variance of X is

$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2 = \int_{-1}^{1} x^2 \frac{x+1}{2} dx - (1/3)^2 = 2/9.$

Example 2. If X has the p.d.f.

$F(x) = 1/x^2$, $1<x,\infty$,

= 0 elsewhere.

Then the mean value of X does not exists, since

$\int_1^{\infty} x \frac{1}{x^2} dx = \lim_{b\to\infty} \int_1^b \frac{1}{x} dx$

$= \lim_{b\to\infty} [\log x]_1^b$

$= \lim_{b\to\infty} (\log b - \log 1)$ does not exist.

Example3. Given that the series

$1/1^2+1/2^2+1/3^2+\ldots\ldots$

converges to $\pi^2/6$. Then

$f(x)=6/\pi^2x^2$, $x=1,2,3,\ldots\ldots$,

= 0 elsewhere,

is the p.d.f. of a discrete type of random variable X. The moment-generating function of this distribution, if it exists, is given by

$M(t)=E(e^{tx})=\sum_x e^{tx} f(x)$

$= \sum_{X=1}^{\infty} 6e^{tx}/\pi^2x^2.$

3.10 CHEBYSHEV'S INEQUALITY:

Theorem :

Let u(X) be a nonnegative function of the random variable X. If E[u(X)] exists, then, for every positive constant c.

$$\Pr[u(X) \geq c] \leq \frac{E[u(X)]}{c}$$

Proof :

The proof is given when the random variable X is of the continuous type; but the proof can be adapted to the discrete case if we replace integrals by sums. Let $A=\{x;u(x) \geq c\}$ and let f(x) denote the p.d.f. of X. Then

$$E[u(X)] = \int_{-\infty}^{\infty} u(x)\, f(x)\, dx = \int_A u(x)\, f(x)\, dx + \int_{A^*} u(x)\, f(x)\, dx.$$

Since each of the integrals in the extreme right-hand member of the preceding equation is nonnegative, the left-hand member is greater than or equal to either of them. In particular,

$$E[u(x)] \geq \int_A u(x)\, f(x)\, dx.$$

However, if $x \in A$, then $u(x) \geq c$; accordingly, the right-hand member of the preceding inequality is not increased if we replace u(x) by c.

Thus

$$E[u(X)] \geq c\int_A f(x)\, dx.$$

Since

$$\int_A f(x)\, dx = \Pr(X \in A) = \Pr[u(X) \geq c],$$

it follows that

$$E[u(X)] \geq c \Pr[u(X) \geq c],$$

Which is the desired result.

Theorem : CHEBYSHEV'S INEQUALITY.

Let the random variable X have a distribution of probability about which we assume that there is a finite variance σ^2. This, of course, implies that there is a mean μ. Then for every $k>0$,

$$\Pr(|X - \mu| \geq k\sigma) \leq 1/k^2,$$

Or equivalently,

$$\Pr(|X-\mu| < k\sigma) \geq 1-1/k^2.$$

Proof. In the above Theorem take $u(X) = (X-\mu)^2$ and $c=k^2\sigma^2$. Then we have

$$\Pr[(X-\mu)^2 \geq k^2\sigma^2] \leq E[X-\mu)^2]/k^2\sigma^2.$$

Since the numerator of the right-hand member of the preceding inequality is σ^2, the inequality may be written

$$\Pr(|X-\mu| \geq k) \leq 1/k^2,$$

Which is the desired result. Naturally, we would take the positive number k to be greater than 1 to have an inequality of interest.

It is seen that the number $1/k^2$ is an upper bound for the probability $\Pr(|X-\mu| \geq k\sigma)$. In the following example this upper bound and the exact value of the probability are compared in special instances.

Example 1:

Let X have the p.d.f.

$$F(x) = 1/2\sqrt{3}, \; -\sqrt{3} < x < \sqrt{3},$$
$$= 0 \text{ elsewhere.}$$

Here $\mu=0$ and $\sigma^2 = 1$. If $k=3/2$, we have the exact probability

$$\Pr(|X-\mu| \geq k\sigma) = \Pr(|X| \geq 3/2) = 1-\int_{-3/2}^{3/2} \tfrac{1}{2}\sqrt{3}\, dx = 1-\sqrt{3}/2.$$

By chebyshev's inequality, the preceding probability has the upper bound $1/k^2 = 4/9$. Since $1-3/2=0.134$, approximately, the exact probability in this case is considerably less than upper bound 4/9. If we take $k=2$, we have the exact probability $\Pr(|X-\mu| \geq 2\sigma) = \Pr(|X|\geq 2)=0$. This again is considerably less than the upper bound $1/k^2=1/4$ provided by Chebyshev's inequality.

In each instance in the preceding example, the probability $\Pr(|X-\mu| \geq k\sigma)$ and its upper bound 1/k2 differ considerably. This suggests that this inequality might be made sharper. However, if we want an inequality that holds for every $k>0$ and holds for all random variables having finite variance, such an improvement is impossible as is shown by the following example.

Example 2.

Let the random variable X of the discrete type have probabilities 1/8,6/8,1/8 at the points x = -1, 0, 1, respectively. Here $\mu=0$ and $\sigma^2=1/4$. If k=2. then $1/k^2=1/4$ and $\Pr(|X-\mu| \geq k\sigma)= \Pr(|X| \geq 1) =1/4$. That is, the probability $\Pr(|X-\mu| \geq k\sigma)$ here attains the upper bound $1/k^2=1/4$. Hence the inequality cannot be improved without further assumptions about the distribution of x^3

Let X be a random variable with mean μ and let E[(X-μ)2k] exist. Show, with d>0, that $\Pr(|X-\mu| \geq d) \leq E[(X-\mu)2k]/d^2k$.

Let X be a random variable such that $\Pr(X\leq 0)=0$ and let $\mu=E(X)$ exist. Show that $\Pr(X\geq 2\mu) \leq 1/2$.

EXERCISE

1. A point is to be chosen in a haphazard fashion from the interior of a fixed circle. Assign a probability p that the point will be inside another circle, which has a radius of one-half the first circle and which lies entirely within the first circle.
2. An unbiased coin is to be tossed twice. Assign a probability P1 to the event that the first toss will be held and that the second toss will be a tail. Assign a probability p_2 to the event that there will be one head and one tail in the two tosses.
3. Find the union $A_1 \cup A_2$ and the intersection $A_1 \cap A_2$ of the two sets A_1 and A_2, where:

(a) $A_1 = \{x;x = 0,1,2\}$, $A_2 = \{x;x = 2,3,4\}$

(b) $A_1 = \{x,o < x<2\}$, $A_2 = \{x;1 \leq x < 3\}$

(c) $A_1 = \{(x,y);0 < x <2,0 <y<2\}$, $A_2 =\{(X,Y);1 <X<3, 1 <Y<3\}$.

4. Find the complement A* of the set A with respect of the space **A** if:

(a) A = 0 {x;0 < x< 1} , A = {x; 5/8 $\leq$ x < 1 }.

(b) **A** = $\{(x,y,z) : x^2+y2 + z2 \leq 1\}$, A = $\{(x,y,z); x^2+y^2+z^2 = 1\}$

(c) **A** = $\{ (x,y) : |x| + |y| \leq 2 \}$, A = $\{(x,y); x^2+y^2<\}$.

5. If the sample space is **C** = $C_1 \cup C_2$ and if $P(C_1) = 0.8$ and $P(C_2) = 0.5$, find $P(C_1 \cap C_2)$

6. Let the sample space be **C** ={ c : 0 < c< ∞ }; Let C ⊂ **C** be defined by C = {c;4 <c < ∞ } and take P(C) = $\int c^{\,c\text{-}x}$ dx. Evaluate P(C) . P(C*), and P(C ∪ C*).

7. Let a card be selected from an ordinary deck of playing cards. The outcome c is one of these 52 cards. Let X(c)= 4 if c is an ace, let X(c)=3 if c is a king, let X(c)=2 if c is a queen, let X(c) = 1 if c is a jack, and let X(c) = 0 otherwise. Suppose that P(C) assigns a probability of 1/52 to each outcome c Describe the induced probability Px(A) on the space A = {x;x = 0,1,2,3,4} of the random variable X.

(8) Let the Space of the random variable X be A = {x;0 < x <1 } If A_1 = {x;0 < x < ½)and A_2 = { x; ½ ≤ x < 1}, find P(A_2) if P(A_1) = ¼ .

(9). Let the space of the random variable X be A = {x;0 < x <10} and let P(A_1) = 3/8 where A_1 = { x;1 < x < 5}.Show that P(A_2) ≤ 5/8, where A_2 = { x;5 ≤ x < 10 }.

(10). Let the subsets A_1 = { x; ¼ < x < ½ } and A_2 = { x; ½ ≤ x < 1 } of the space A = {x;0 < x , 1} of the random variable X be such that P(A_1) = 1/8 and P(A_2) = ½. Find P($A_1 \cup A_2$), P(A^*_1), and P($A^*_1 \cap A^*_2$)

(11) Let A_1 = {(x,y); x ≤ 2, y≤4, A_2 = {(x,y); x≤ 2, y ≤ 1} A_3 = {(x,y); x ≤ 0, y ≤ 4}, and A_4 = {(x,y); x ≤ 0, y ≤ 1} be subsets of the space A of two random variables X and Y, which is the entire two-dimensional plane. If P(A_1) = 7/8, P(A_2) = 4/8 P(A_3) = 3/8 and P(A_4) = 2/8, find P(A_5), where A_5 = {(x,y); 0 < x ≤ 2, 1 < y ≤ 4}.

(12) Give $\int_A$ [1/π (1+x^2)] dx, where A ⊂ A = { x; -∞ < x < ∞ } show that the integral could serve as a probability set function of a random variable X whose space is A

(13). For each of the following, find the constant c so that f(x) satisfies the conditions of being a p.d.f. of one random variable X.

(a) f(x)=c(2/3)x,x=1,2,3,.... , zero elsewhere.

(b) f(x)=cxe^{-x},0<x<∞,zero elsewhere.

(14) Let f(x)=x/15, x=1,2,3,4,5, zero elsewhere, be the p.d.f. of X. Find pr (X= 1 or 2), Pr(1/2<X<5/2), and Pr(1≤ X ≤2).

(15). Show that $\int_0^\infty xe^{-x}\,dx = \int_0^\infty e^{-x}\,dx = 1$,

and, for k≥1, that (by integrating by parts)

$\int_0^\infty x^k e^{-x}\,dx = k\int_0^\infty x^{k-1}e^{-x}\,dx$.

(a) What is the value of $\int_0^\infty$ xne^{-x} dx, where n is a nonnegative integer?

(16) Given the distribution function

$F(x) = 0, x<-1,$

$= x+2/4, -1\le x<1,$

$= 1, 1\le x.$

Sketch the graph of F(x) and then compute: (a) $Pr(-1/2 < X \le ½)$; (b) $Pr(X=0)$; (c) $Pr(X=1)$;

(d) $Pr(2<X \le 3)$.

(17) Let $f(x)= (4-x)/16, -2 < x < 2$, zero elsewhere, be the p.d.f. of X.

(a) Sketch the distribution function and the p.d.f. of X on the same set of axes.

(b) If $Y = |X|$, compute $Pr(Y\le1)$.

(c) If $Z = X^2$, compute $Pr(Z\le1/4)$.

Let F(x) be the distribution function of the random variable X. If m is a number such that $F(m) = ½$, show that m is a median of the distribution.

(18) Compute the probability of being dealt at random and without replacement a 13-card bridge hand consisting of: (a) 6 spades, 4 hearts, 2 diamonds, and 1 club ; (b) 13 cards of the same suit.

(19) Three distinct integers are chosen at random from the first 20 positive integers. Compute the probability that; (a) their sum is even; (b) the product is even.

(20) Let X have the uniform distribution given by the p.d.f. $f(x) =1/5$, $x = -2,-1,0,1,2$, zero elsewhere. (a) Find the p.d.f. of $Y=X^2$.

(21) Let X have the p.d.f. $f(x)=(x+2)/18, -2<x<4$, zero elsewhere. Find $E(X)$, $E[(X+2)^3]$, and $E(6X-2(X+2)^3]$.

(22) Let the p.d.f. of X and Y be $f(x,y) = e^{-x-y}$, $0<x<\infty$, $0<y<\infty$, zero elsewhere. Let $u(X,Y) =X, v(X,Y) =Y$ and $w(X,Y)=XY$. Show that $E[u(X,Y)]. E[v(X,Y)] = E[w(X,Y)]$.

(23). Let X have a p.d.f. f(x) that is positive at x= -1,0,1 and is zero elsewhere.

(a) If f(0)=1/2, find $E(X^2)$.

(b) If f(0)=1/2 and if E(X)=1/6, determine f(-1) and f(1).

(24) Find the mean and varience, if they exist, of each of the following distributions.

(a) $f(x)=3!/x!(3-x)!\ (1/2)^3$, x = 0,1,2,3, zero elsewhere.

(b) $f(x)=6x(1-x)$, $0<x<1$, zero elsewhere.

(c) $f(x)= 2/x^3$, $1 < x < \infty$, zero elsewhere.

(25) Let $f(x)=(1/2)^3$, x= 1, 2, 3, ..., zero elsewhere, be the p.d.f. of the random variable X. Find the moment-generating function, the mean, and the variance of X.

(26) For each of the following probability density functions, compute $Pr(\mu - 2\sigma < X < \mu + 2\sigma)$.

(a) $f(x)= 6x(1-x)$, $0 < x < 1$, zero elsewhere.

(b) $f(x)= (1/2)x$, x= 1, 2, 3, ..., zero elsewhere.

(27) Let the random variable X have the p.d.f.

$F(x) = p,\quad x = -1,1,$

$= 1 - 2p,\ x=0,$

$= 0$ elsewhere,

where $0 < p < ½$. Find the measure of kurtosis as a function of p. Determine its value when p=1/3, p=1/5, p=1/10, and p=1/100. Note that the kurtosis increases as p decreases.

UNIT IV

CONDITIONAL PROBABILITY AND STOCHASTIC INDEPENDENCE

TABLE OF CONTENTS

4.1 CONDITIONAL PROBABILITY

Let the probability set function P(C) be defined on the sample space and let C_1 be a subset of such that $P(C_1)>0$.The conditional probability of the event C_2, relative to the event C_1: or, more briefly, the conditional probability of C_2, given C_1 is denoted hj $p(c_2/c_1)$ and is defined hj

$$P(C_1/C_1)=1 \text{ and } P(C_2/C_1)= P(C_1 \cap C_2/C_1).$$

Hence

$$P(C_2/C_1)= P(C_1 \cap C_2)/P(C_1)$$

Is a suitable definition of the conditional probability of the event C_2, given the event C_1, provided $P(C_1) > 0$.

Let P denote the probability set function of the induced probability on A. If A_1 and A_2 are subsets of A, the conditional probability of the event A_2, given the event A_1, is

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

Provided $P(A_1)>0$.

Example. A hand of 5 cards is to be dealt at random and without replacement from an ordinary deck of 52 playing cards. The conditional probability of an all-spade hand (C_2), relative to the hypothesis that there are at least 4 spades in the hand (C_1), is, since $C_1 \cap C_2=C_2$,

$$P(C_2/C_1)=P(C_2)/P(C_1) = \frac{{}^{13}C_5 / {}^{52}C_5}{[{}^{13}C_4 \times {}^{39}C_1 + {}^{13}C_5 / {}^{52}C_5]}$$

Example

A bowl contains eight chips. Three of the chips are red and the remaining five are blue. Two chips are to be drawn successively, at random and without replacement. We want to compute the probability that the first draw results in a red chip (C_1) and that the second draw results in a blue chip(C_2). It is reasonable to assign the following probabilities:

$P(C_1) = 3/8$ and $P(C_2/C_1) = 5/7$.

Thus, under these assignments, we have $P(C_1 \cap C_2) = 3C_8 \times 5C_7 = \frac{15}{56}$

Example

From an ordinary deck of playing cards, cards to be dawn successively, at random and without replacement. The probability that the third spade appears on the sixth draw is computed as follows. Let C_1 be the event of two spades in the first five draws and let C_2 be the event of a spade on the sixth draw. Thus the probability that we wish to compute is $P(C1 \cap C2)$. It is reasonable to take

$$P(C_1) = \frac{13C_2 \times 30C_3}{32C_5} \quad \text{and} \quad P(C_2/C_1) = 11/47.$$

The desired probability $P(C_1 \cap C_2)$ is then the product of these two numbers. More generally, if X+3 is the number of draws necessary to produce exactly three spades, a reasonable probability model for the random variable X is given by the p.d.f.

$$F(x) = \frac{13C_2 \times 39C_x}{52C_{2+x)}} \quad \frac{11,}{50\text{-}x} \qquad x = 0, 1,2,\ldots\ldots 39$$

= 0 elsewhere.

Then the particular probability which we computed is

$P(C_1 \cap C2) = Pr(X=3) = f(3)$.

4.2 MARGINAL AND CONDITIONAL DISTRIBUTIONS:

Let $f(x_1,x_2)$ be the p.d.f. of two random variables X_1 and X_2. $F(x_1,x_2)$ is the joint p.d.f. of the random variables X_1 and X_2. Consider the event $a<X_1<b$, $a<b$. This event can occur when and only when the event $a<X_1<b$, $-\infty<X_2<\infty$ occurs; that is, the two events are equivalent, so that they have the same probability. But the probability of the latter event has been defined and is given by

$Pr(a<X_1<b,-\infty<X_2<\infty) = \int_a^b \int_{-\infty}^{\infty} f(x_1,x_2)dx_2\, dx_1$

for the continuous case, and by

$Pr(a<X_1<b,-\infty<X_2<\infty)=\sum_{a<x_1<b}\sum_{x_2} f(x_1,x_2)$

for the discrete case.

Again, $f_2(x_2) = \int_{-\infty}^{\infty} f(x_1,x_2)dx_1$ (continuous case),

$= \sum f(x_1,x_2)$ (discreate case)

is called the marginal p.d.f. of X_2 Where $f_2(x_2)$ is the p.d.f of x_2 alone

Example Let the joint p.d.f. of X_1 and X_2 be

$f(x_1,x_2) = \frac{x_1+x_2}{21}$, $x_1=1,2,3$, $x_2 = 1,2$

= 0 elsewhere

Then

$Pr(X_1=3)=f(3,1)+f(3,2)=3/7$

and $Pr(X_2=2)=f(1,2)+f(2,2)+f(3,2)= 4/7$.

On the other hand the marginal p.d.f of X_1 is

$$f_1(X_1) = \sum_{x_2=1}^{2} \frac{x_1+x_2}{21} = \frac{2x_1+3}{21}, \quad x_1=1,2,3$$

zero elsewhere, and the marginal p.d.f of x_2 is

$$f_2(x_2) = \sum \frac{x_1+x_2}{21} = \frac{6+3x_2}{21}, \quad x_2=1,2$$

zero elsewhere. Thus the preceding probabilities may be computed as $Pr(X_1=3) = f_1(3)=3/7$ and $Pr(X_2=2)=f_2(2)=4/7$

Example Let X_1 and X_2 have the joint p.d.f

$f(x_1,x_2) = 2,\ 0<x_1<x_2<1$

= 0 elsewhere

Then the marginal probability density functions are respectively,

$f_1(x_1) = {}_{x_1=1}\int^{2} 2dx_2 = 2(1=x_1), 0<x_1<1$

$= 0$ elsewhere

and $f_2(x_2)_0\int^{x2} 2dx_1 = 2x_2 ,0<x_2<1$

$= 0$ elsewhere

The conditional p.d.f of X_1 given $X_2 = x_2$, is

$f(x_1/x_2) = 2/2X_2 = 1/x_2, 0<x_1<x_2, 0<x_2<1$

$= 0$ elsewhere

Here the conditional mean and conditional variance of X_1, given $X_2 = x_2$ are, respectively,

$$E(X_1|x_2) = {}_{-\infty}\int^{\infty} x_1 f(x1|x2)dx_1$$

$$= {}_0\int^{x_2} x_1 \cdot 1/x^2\, dx_2$$

$$= \frac{x^2}{2}, \quad 0<x_2<1,$$

and $E[(X_2\text{-}E(X_1|x_2)]^2 /x_2\} = {}_0\int^{x2} (x_1\text{-}x_2/_2)^2 (1/x_2)dx_1$

$$= x_2^2/12, 0<x_2<1.$$

Finally, we shall compare the values of Pr $(0<X_1<1/2|X_2 = 3/4)$ and $\Pr(0<X_1<1/2)$. We have

$\Pr(0<X_1<1/2|X_2 = 3/4) = \int_0^{1/2} f(x_1|3/4)dx_1 = \int_0^{1/2} (4/3)dx_1 = 2/3$

but

$\Pr(0<X1<1/2 = \int_0^{1/2} f_1(x_1)\, dx_1 = \int_0^{1/2} 2(1-x_1)dx_1 = 3/4)$

Let the random variables, $X_1, X_2, X_3 .... X_n$ have the joint p.d.f $(x_1, x_2, x_3 ..... x_n)$. If the random variable are of the continuous type, then by an argument similar to the two - variable case, we have for every a<b, $\Pr(a < X_1 < b) = {}_a\int^b f_1(x_1)dx_1$

Where $f_1(x_1)$ is defined by the (n-1) fold integral

$f_1(x_1) = {}^{\infty}\int_{-\infty} \ldots {}^{\infty}\int_{-\infty} f(x_1, x_2, \ldots\ldots\ldots\ldots xn)\, dx_2 \ldots\ldots dxn$

Accordingly $f_1(x_1)$ is the p.d.f of the one random vaiable X_1 and $f_1(x_1)$ is called the marginal p.d.f of X_1. The marginal probability density functions, $f_2(x_2)$, ... $f_4(x_4)$ of $x_2, \ldots\ldots x_n$ respectively are similar (n-1) fold integrals. Each marginal p.d.f has been a p.d.f of one random variable. It is convenient to extend this terminology to joint probability density functions. Let f(x1,x2,.....xn) be the join p.d.f of the n random variables X_1, X_2, ... X_n) Take any group of k <n of these random variables and let us find the join p.d.f of them. This joint p.d.f. is called the marginal p.d.f. of this particular group of k variables. The marginal p.d.f. of X_2, X_4, X_5 is the joint p.d.f. of this particular group of three variables, namely,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1,x_2,x_3,x_4,x_5,x_6)\, dx_1\, dx_3\, dx_6$$

if the random variables are of the continuous type.

If $f_1(x_1) > 0$, the symbol $f(x_2......,x_n \mid x_n) = \dfrac{f(x_1,x_2 ....x_n)}{f_1(x_1)}$

and $f(x_2.......x_n \mid x_1)$ is called the joint conditional p.d.f. of X_2......X_n given $X_1 = x_1$. The joint conditional p.d.f. of any n-1 random variables, say X_1,........X_{i-1} , X_{i+1} , ... $X_{n'}$ given $X_i = x_i$ is defined as the joint p.d.f. of X_1, X_2, ... X_n divided by marginal p.d.f. $f_i(x_i)$, provided $f_i(x_i) > 0$. More generally, the joint conditional p.d.f. of n-k of the random variables, for given values of the remaining k variables, is defined as the joint p.d.f. of the n variables divided by the mariginal p.d.f. of the particular group of k variables, provided the latter p.d.f is positive.

The conditional expectations of $u(X_2,......,X_n)$ given $X_1=x_1$,is , for random variables of the continuous type , given by $E[u(X_2,...,X_n)|x_1]$

$$= \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} u(x_2,....x_n) f(x_2,...x_n|x_1)\, dx_2....dx_n$$

provided $f_1(x_1)>0$ and the integeral converges(absolutely).

4.3 THE CORRELATION COEFFICIENT

Let X,Y ,and Z denote random variables that have joint p.d.f. f(x,y,z). The means of X,Y, and Z , say μ_1 ,μ_2 and μ_3, are obtained by taking u(x,y,z) to be x,y, and z, respectively; and the variances of X,Y and Z, say σ_1^2 ,σ_2^2 and σ_3^2, are obtained by setting the function u(x,y,z)equal to $(x-\mu_1)^2$,$(y-\mu_2)^2$, and $(z-\mu_3)^2$, respectively.

$$E[(X-\mu_1)(Y-\mu_2) = E(XY-\mu_2X-\mu_1Y+\mu_1\mu_2)$$
$$= E(XY)-\mu_2E(X)-\mu_1E(Y)+\mu_1\mu_2$$
$$= E(XY)-\mu_1\mu_2.$$

This number is called the covariance of X and Y. The covariance of X an Z is given by $E[(X-\mu_1)(Z-\mu_3)]$, an the covariance of Y an Z is $E[(Y-\mu_2)(Z-\mu_3)]$.

If each of σ_1 and σ_2 is positive , the number

$$\rho_{12} = \frac{E[(X-\mu_1)(Y-\mu_2)]}{\sigma_1\sigma_2}$$

is called the correlation coefficient of X and Y.

Example Let the random variable X and Y have the joint p.d.f.

$F(x,y)=x+y, \quad 0<x<1, 0<y<1,$

$=0$ elsewhere,

Compute the correlation coefficient of X and Y. When only two variables are under consideration , we shall denote the correlation coefficient by ρ. Now

$\mu_1=E(X) = \int_0^1 \int_0^1 x(x+y)dxdy=7/12$

and

$\sigma_1^2=E(x_2)-\mu_1^2=\int_0^1 \int_0^1 x_2(x+y)dxdy-(7/12)^2=11/144$

Similarly, $\mu_2=E(y)=7/12$ and $\sigma^2=E(y_2)-\mu_2^2=11/44$

The covarience of X and Y is

$e(XY)-\mu_1\mu_2 = \int_0^1\int_0^1 xy(x+y)dxdy-(7/12)_2=-1/144.$

Accordingly, the correlation coefficient of X and Y is

$\rho = -1/144$

$(11/144)\ (11/144) = -1/11.$

Example

Let the continuous type random variables X and Y have the joint p.d.f

$F(x,y)=e^{-y}, \quad 0<x<y<\infty$

$=0$ elsewhere

The moment generating function of this joint distribution is

$M(t1,t2 = \int_0^\infty \int_x^\infty \exp(t_1x+t_2y-y)dydx$

$= \dfrac{1}{(1-t_1-t_2)(1-t_2)}$

provided $t_1+t_2<1$ and $t_2<1$. For this distribution, Equations

$$\sigma_1^2 = E(x^2) - \mu_1^2 = \frac{\partial^2 M(0,0)}{\partial_{t1}^2} - \mu_1^2$$

becomes $\mu_1=1, \quad \mu_2=2$

$\sigma_1^2=1 \quad \sigma_2^2=2$

$E(X-\mu_1)(Y-\mu_2)]=1$

Further more, the moment-generating functions of the marginal distributions of X and Y are, respectively.

$M(t_1,0) = \frac{1}{1-t_1}$, $t_1<1$

$M(0,t2) = \frac{1}{(1-t_2)^2}$, $t_2<1$

These moment-generating functions are, of course, respectively, those of the marginal probability density functions,

$f_1(x) = \int_x^{\infty} e^{-y} = e^{-x}$, $0<x<\infty$

zero elsewhere, and

$f_2(y)=e^{-y} \int_x^{\infty} dx=ye^{-y}$, $0<y<\infty$

Zero elsewhere.

4.4 STOCHASTIC INDEPENDENCE

Let X_1 and X_2 denote random variables of either the continuous or the discrete type which have the joint p.d.f. $f(x_1,x_2)$ and marginal probability density functions $f_1(x_1)$ and $f_2(x_2)$, respectively.

The joint p.d.f. $f(x_1,x_2)$ as

$$f(x_1,x_2) = f(x_2 \mid x_1) f_1(x_1).$$

Definition

Let the random variables X_1 and X_2 have the joint p.d.f. $f(x_1,x_2)$ and the marginal probability density functions $f_1(x1)$ and $f_2(x_2)$ respectively. The random variables X_1 and X_2 are said to be stochastically independent if, and only if, $f(x1,x2) \equiv f_1(x_1) f_2(x_2)$. Random variables that are not stochastically independent are said to be stochastically dependent.

Example : Let the joint p.d.f. of X and X_2 be

$$f(x_1,x_2) = x_1 + x_2 ,\ 0 < x_1 < 1, 0 < x_2 < 1.$$
$$= 0 \text{ elsewhere.}$$

It will be shown that X_1 and X_2 are stochastically dependent. Here the marginal probability density functions are

$f_1(x_1) = \int_{-\infty}^{\infty} f(x_1,x_2)\, dx_2 = \int_0^2 (x_1+x_2)\, dx_2 = x_1 + ½$, $0 < x_1 < 1$,

$= 0$ elsewhere

$f_2(x_2) = \int_{-\infty}^{\infty} f(x_1,x_2)\, dx_1 = \int_0^1 (x_1+x_2)\, dx_1 = ½ + x2$, $0 < x_2 < 1$,

$= 0$ elsewhere

Since $f(X_1,X_2) \not\equiv f_1(X_1)f_2(X_2)$, the random variable X_1 and X_2 are stochastically dependent.

The following theorem makes it possible to assert, without computing the marginal probability density functions, that the random variables X_1 and X_2 of Example above are stochastically dependent.

Theorem (1)

Let the random variables X_1 and X_2 have the joint p.d.f $f(x_1,x_2)$. Then X_1 and X_2 are stochastically independent if and only if $f(x_1,x_2)$ can be written as a product of a non negative function of x_1, alone and a non negative function of x_2 alone. That is,

$$F(x_1,x_2) = g(x_1)h(x_2),$$

Where $g(x_1)>0$, $x_1 \in A_1$, zero elsewhere, and $h(x_2)>0$, $x_2 \in A_2$, zero elsewhere.

Proof.

If X_1 and X_2 are stochastically independent, then $f(x_1,x_2) \equiv f_1(x_1)f_2(x_2)$, where $f_1(x_1)$ and $f_2(x_2)$ are the marginal probability density functions of X_1 and X_2, respectively. Thus, the condition $f(x_1,x_2) \equiv g(x_1)h(x_2)$ is fulfilled.

Conversely, if $f(x_1,x_2) \equiv g(x_1)h(x_2)$, then, for random variables of the continuous type, we have

$$f_1(x_1) = \int_{-\infty}^{\infty} g(x_1)h(x_2)dx_2 = g(x_1) \int_{-\infty}^{\infty} h(x_2)dx_2 = c_1 g(x_1)$$

and

$$f_2(x_2) = \int_{-\infty}^{\infty} g(x_1)h(x_2)dx_1 = h(x_2) \int_{-\infty}^{\infty} g(x_1)dx_1 = c_2 h(x_2),$$

where c_1 and c_2 are constants, not functions of x_1 or x_2. Moreover $c_1c_2 = 1$ because

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x_1)h(x_2)dx_1dx_2 = \left[\int_{-\infty}^{\infty} g(x_1)dx_1\right] \left[\int_{-\infty}^{\infty} h(x_2)dx_2\right] = c_2c_1.$$

These results imply that

$$f(x_1,x_2) \equiv g(x_1)h(x_2) \equiv c_1g(x_1)c_2h(x_2) \equiv f_1(x_1)f_2(x_2).$$

Accordingly, x_1 and x_2 are stochastically independent.

From the above example we see that the joint p.d.f.

$$f(x_1,x_2) = x_1+x_2,\ 0<x_1<1, 0<x_2<1, =0 \text{ elsewhere},$$

cannot be written as the product of a nonnegative function of x_1 alone and a nonnegative function of x_2 alone. Accordingly, X_1 and X_2 are stochastically dependent.

Theorem 2:- If X_1 and X_2 are stochastically independent random variables with marginal probability density functions $f_1(x_1)$ and $f_2(x_2)$, respectively, then

$Pr(a<X_1<b, c<X_2<d) = Pr(a<X_1<b)\ Pr(c<X_2<d)$ for every $a<b$ and $c<d$, where a,b,c, and d are constants.

Proof. From the stochastic independence of X_1 and X_2, the joint p.d.f. of X_1 and X_2 is $f_1(x_1)f_2(x_2)$. Accordingly, in the continuous case,

$$Pr(a<X_1<b, c<X_2<d) = \int_a^b \int_c^d f_1(x_1)f_2(x_2)dx_2\, dx_1$$

$$=[\int_a^b f_1(x_1)dx_1][\int_c^d f_2(x_2)dx_2]$$

$$=Pr(a<X_1<b)Pr(c<X_2<d);$$

or, in the discrete case,

$$Pr(a<X_1<b, c<X_2<d) = \sum_{a<x_1<b} \sum_{c<x_2<d} f_1(x_1)f_2(x_2)$$

$$= [\sum_{a<x_1<b} f_1(x_1)] [\sum_{c<x_2<d} f_2(x_2)]$$

$$= Pr(a<X_1<b)Pr(c<X_2<d),$$

Example

In first Example X_1 and X_2 were found to be stochastically dependent. There, in general,

$$Pr(a<X_1<b, c<X_2<d) \neq Pr(a<X_1<b)Pr(c<X_2<d).$$

For instance,

$$Pr(0<X_1<1/2, 0<X_2<1/2) = \int_0^{1/2} \int_0^{1/2} (x_1+x_2)dx_1 dx_2 = 1/8,$$

whereas

$$Pr(0<X_1<1/2) = \int_0^{1/2} (x_1+1/2)dx_1=3/8$$

and $Pr(0<X_2<1/2) = \int_0^{1/2} (1/2+x_2)dx_2 = 3/8$

Theorem 3. Let X_1 and X_2 denote random variables that have the joint p.d.f. $f(x_1,x_2)$ and the marginal probability density functions $f_1(x_1)$and $f_2(x_2)$, respectively. Furthermore, let $M(t_1,t_2)$ denote the moment-generating function of the distribution. Then X_1 and X_2 are stochastically independent if and only if $M(t_1,t_2) = M(t_1,0)M(0,t_2)$.

Proof. If X_1 and X_2 are stochastically independent, then

$$M(t_1,t_2) = E(e^{t_1x_1+t_2x_2})$$

$$= E(e^{t_1x_1}e^{t_2x_2})$$

$$= E(e^{t_1x_1})E(e^{t_2x_2})$$

$$= M(t_1,0)M(0,t_2).$$

Thus the stochastic independence of X_1 and X_2 implies that the moment-generating function of the joint distribution factors into the product of the moment-generating functions of the two marginal distributions.

Suppose that the moment-generating function of the joint distribution of X_1 and X_2 is given by $M(t_1,t_2)=M(t_1,0)M(0,t_2)$. Now X_1 has the unique moment-generating function which, in the continuous case, is given by

$$M(t_1,0) = \int_{-\infty}^{\infty} e^{t_1x_1} f_1(x_1)dx_1).$$

Similarly, the unique moment-generating function of X_2, in the continuous case, is given by

$$M(0,t_2) = \int_{-\infty}^{\infty} e^{t_2x_2} f_2(x_2)dx_2.$$

Thus we have

$$M(t_1,0)M(0,t_2)=[\int_{-\infty}^{\infty} e^{t_1x_1} f_1(x_1)dx_1][\int_{-\infty}^{\infty} e^{t_2x_2} f_2(x_2)dx_2]$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x_1+t_2x_2} f_1(x_1)f_2(x_2)dx_1\, dx_2.$$

We are given that $M(t_1,t_2)=M(t_1,0)M(0,t_2)$:so

$$M(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x_1+t_2x_2} f_1(x_1)f_2(x_2)dx_1\, dx_2$$

But $M(t_1,t_2)$ is the moment-generating function of X_1 and X_2. Thus also

$$M(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x_1+t_2x_2} f(x_1,x_2)dx_1dx_2.$$

The uniqueness of the moment generating function implies that the two distributions of probability that are described by $f_1(x_1)f_2(x_2)$ and $f(x_1,x_2)$ are the same. Thus

$f(x_1,x_2) \equiv f_1(x_1)f(x_2)$

That is, if $M(t_1,t_2)=M(t_1,0)M(0,t_2)$, then X_1 and X_2 are stochastically independent.

Some Special Distributions

4.5 The Binomial, Trinomial and Multinomial Distribution:

If n is a positive integer, that $(a+b)^n = \sum_{x=0}^{n} {}^nC_x\, b^x a^{n-x}$; x=0,1,......n

Consider the function defined by

$$f(x) = nC_x\, p^x (1-p)^{1-x}, \quad x = 0,1,2,\ldots\ldots,n,$$
$$= 0 \text{ elsewhere},$$

where n is a positve integer and $0<p<1$,. Under these conditions it is clear that $f(x)\geq 0$ and that

$$\sum_x f(x) = \sum_{x=0}^{n} nC_x\ nC_x\, p^x(1-p)^{n-x}$$

$$= [(1-p)+p]^n = 1.$$

That is f(x) satisfies the conditions of being a p.d.f of a random variable X of the discrete type. A random variable X that has a p.d.f. of the form of f(x)is said to have a binominal distribution, and any such f(x)is called a binominal p.d.f. A binomial distribution will be denoted by the symbol b(n,p).

If we say that X is b(5,1/3),we mean that X has the binomial p.d.f.

$$f(x) = 5C_x \, 1C_3 \, 2C_3^{5-x}, \quad x=0,1........,5,$$

$$= 0 \text{ elsewhere.}$$

Example 1. The binomial distribution with p.d.f.

$$f(x) = 7C_x \, 1C_2^x \left(\frac{1-1}{2}\right)^{n-x} \quad x = 0,1,2..........,7,$$

$$= 0 \text{ else where}$$

has the moment generating function

$$M(t)=(1/2+1/2^{et})^7,$$

has mean $\mu = np = 7/2$, and has variance $\sigma^2 = np(1-p)=7/4$. Furthermore, if X is the random vaiable with this distribution, we have

$Pr(0\leq 1) = \sum f(x)= 1/128+7/128=8/128$ and

$Pr(X=5)=f(5)$

$= (7!/5!2!)(1/2)^5(1/2)^2 = 21/128$

Example 2. If the moment generating function of a random variable X is

$M(t)=(2/3+1/3\ ^{et})^5$,

then X has a binomial distribution with n=5 and p=1/3; that is the p.d.f od X is

$$f(x) = 5C_x \, 1C_3^{\,x} \, 2C_3^{5-x}, \; x = 0,1,2.....5$$

$$= 0 \text{ elsewhere}$$

Here $\mu = np = 5/3$ and $\sigma^2 = np(1-p) =10/9$

Example 3

Consider a sequence of independent repetition of a random experiment with constant probability p of success. Let the random variable Y denote the total number of failures in the sequence before the rth success that is, Y+r is equal to the number of trials necessary to produce exactly r success. here r is a fixed positive integer. To determine the p.d.f of Y, Let y be an element of {y;y=0,1,2,.....}. Then, by the multiplication rule of probabilities, Pr (Y=y) = g(y) is equal to the product of the probability

$$(y+r-1)\, C_{r-1}\, p^{r-1}(1-p)^r$$

of obtaining exactly r-1 success in the first y+r-1 trials and the probability p of a success on the $(y+r)^{th}$ trial. Thus the p.d.f g(y) of Y is given by

$g(y) = y+r-1\ C_{r-1}\ p^r(1-p)^y, \ y = 0,1,2.....$

$= 0$ elsewhere

A distribution with a p.d.f. of the form g(y) is called a negative binomial distribution; and any such g(y) is called a negative binomial p.d.f. The distribution derives its name from the fact that g(y) is a general term in the expansion of $pr[1-(1-p)]^{-r}$. It is left as an exercise to show that the moment generating function of this distribution is $M(t) = p^r\ [1-(1-(1-p)e^t)^{-r}$ for $t<-\ln(1-p)$. If r=1, then Y has the p.d.f.

$g(y) = p(1-p)^y, \ y = 0,1,2......$

zero elsewhere, and the moment generation function $M(t)= p[1-(1-p)^{et}]-1$ In this special case, r =1, we say that Y has a geometric distribution

4.6 THE POISSION DISTRIBUTION

The series $1+m+m^2/2!+m^3/3!+.......= \sum_{n=0}^{\alpha} m^x/x!$

converge, for all values of m, to e^m. Consider the function $f(x)$ defined by

$f(x) = m^x e^{-m}/x!, \ x = 0,1,2.....$

$= 0$ elsewhere,

where m>0. Since m>0, then $f(x) \geq 0$ and

that is $f(x)$ satisfies the conditions of being a p.d.f of a discrete type of random variable. A random variable that has a p.d.f of the form $f(x)$ is said to have a poisson distribution, and any such $f(x)$ is called a poisson p.d.f.

Example 1. Suppose that X has a poisson distribution with $\mu=2$. Then the p.d.f od X is

$f(x) = 2xe\text{-}2/x!, \ x = 0,1,2....$

$= 0$ elsewhere

The variance of this distribution is $\sigma^2=\mu=2$. If we wish to compute $Pr(1\leq X)$, we have

$Pr(1\leq X) = 1-Pr(X=0)$

$= 1-f(0)=1-e^{-2}=0.865$

approximately.

Example

If the moment generating function of a random vriable X is

$$M(t) = e^{4(et-1)}$$

then X has a poisson distribution with $\mu=4$. Accordingly, by way of example,

$$\Pr(X=3) = 4^3e^{-4/3}! = \frac{32e^{-4}}{3}$$

(or) $\Pr(X = 3) = \Pr(X \leq 3) - \Pr(X \leq 2) = 0.433 - 0.238 = 0.195$

4.7. THE GAMMA AND CHI-SQUARE DISTRIBUTIONS

The Gamma function of X is

$$\Gamma(\alpha) = {}_0\int^{\infty} y^{\alpha-1} e^{-y}\, dy$$

If $\alpha = 1$, Clearly

$$\Gamma(1) = {}_0\int^{\infty} e^{-y}\, dy = 1$$

If $\alpha > 1$, an integration by parts shows that

$$\Gamma(\alpha) = (\alpha - 1)\, {}_0\int^{\infty} y^{a-2} e^{-y}\, dy = (\alpha - 1)\Gamma(\alpha-1)$$

Accordingly, if α is a positive integer greater than 1,

$$\Gamma(\alpha) = (\alpha-1)(\alpha-2) \ldots\ldots (3)(2)(1)\Gamma(1) = (\alpha-1)!$$

since $\Gamma(1) = 1$

In the integral that defines $\Gamma(a)$, let us introduce a new variable x by writing $y = x/\beta$, where $\beta >$ 0 Then.

$$\Gamma(\alpha) = {}_0\int^{\infty} \left(\frac{x}{\beta}\right)^{a-1} e^{-n/\beta} \left(\frac{1}{\beta}\right) dx$$

or, equivalently,

$$1 = {}_0\int^{\infty} 1/\Gamma(\alpha)\beta^{\alpha}\ x^{a-1}e^{-x/\beta}\ dx,$$

Since $\alpha>0$, $\beta>0$, and $\Gamma(\alpha)>0$, we see that

$$f(x) = 1/\Gamma(\alpha)\beta^{\alpha} x^{\alpha-1} e^{-x/\beta}, 0<x<\infty,$$
$$= 0 \text{ else where}$$

is a p.d.f. of a random variable of the continuous type.

Example

Let X be a random variable such that

$$E(X^{m)} = \frac{(m+3)!}{3!} 3^m, m=1,2,3 \ldots\ldots$$

Then the moment generating function of X is given by the series

$$M(t) = 1 + \frac{4!3}{3!1!}t + \frac{5!3^2}{3!2!}t^2 + \frac{6!3^3}{3!3!}t^3 + \ldots$$

This, however is the Maclaurin's series for $(1-3t)^{-4}$ provided that $-1<3t<1$. Accordingly, X has a gamma distribution with $\alpha=4$ and $\beta=3$

Example

If X has the moment generating function $M(t)= 1-2t)^{-8}$, $t<1/2$ then X is x^2 (16)

If the random variable X is $x^2(r)$, then with $c_1 \le c_2$, we have

$Pr(c_1 \le X \le c_2) = Pr(X \le c_2) - Pr(X \le c_1)$,

since $Pr(X=c_1)=0$. To compute such a probability , we need the value of an integral like

$$Pr(X \le x) = \int_0^x 1/\Gamma(r/2)2^{r/2}\ \omega^{r/2-1}\ e^{-w/2}\,dw$$

Example

Let X have a gamma distribution with $\alpha=r/2$, where r is a positive integer , and $\beta>0$. define the random variables $Y = 2X/\beta$. We seek the p.d.f of Y. Now the distribution function of Y is

$$G(y) = Pr(Y<y) = Pr(X< \beta y/2)$$

If $y < 0$, then $G(y)=0$; but if $y>0$ then

$$G(y) = \int_0^{\beta y/2} 1/\Gamma\beta r/2\ \ \beta y/2\ \ )r/2-1\ \ e\ \ -y/2$$

$$= 1/\Gamma(r/2)\ 2\ r/2\ y\ r/2-1\ e\ -y/2$$

if $y>0$. That is Y is $x2(r)$

4.8 THE NORMAL DISTRIBUTION

Consider the integral

$$I = \int_{-\infty}^{\infty} \exp(-y^2/2)dy.$$

This integral exits because the integrand is a positive continuous function which is bounded by an integrable function; that is,

$0<\exp(-y^2/2)<\exp(-|y|+1)$, $-\infty<y<\infty$, and

$$\int_{-\infty}^{\infty} \exp(-|y|+1)dy = 2e$$

To evaluate the integral I, we note that $I>0$ and that I^2 may be written

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\frac{(-y^2+z^2)}{2}\,dydz$$

Example 1

If X has the moment generating function

$$M(t) = e^{2t+32t2}$$

then X has a normal distribution with $\mu=2,\sigma$). Thus, if we say hat the random variable X is n(0,1), we mean that X has a normal distribution with mean $\mu=0$ and variance $\sigma^2=1$, so that the p.d.f of X is

$$f(x) = 1/\sqrt{2\pi}\, e^{-x2/2}, \ -\infty < x\ \infty$$

If we say that X is n(5,4), we mean that X has a normal distribution with mean $\mu=5$ and variance $\sigma^2=4$, so that the p.d.f of X is

$$f(x) = \frac{1}{2\sqrt{2\pi}} \exp\left[\frac{(x-5)^2}{2(4)}\right] \quad -\alpha<x<\alpha$$

Moreover, if

$$M(t) = e^{t2/2,}$$

then X is n(0,1)

The graph of

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp -\frac{(x-\mu)^2}{2\sigma^2}, \quad -\alpha<x<\alpha,$$

is seen (1) to be symmetric about a vertical axis through $s=\mu$ and (3) to have the x-axis as a horizontal asymptote. It should be verified that (4) there are points of inflection at $x = \mu \pm \sigma$.

Theorem 1.

If the random variable X is $n(\mu, \sigma^2)$, $\sigma^2 > 0$, then the random variable $W = (X - \mu) / \sigma$ is n (0,1).

Proof. :- The distribution function $G(\omega)$ of w is, since $\sigma > 0$,

$$G(\omega) = Pr(X - \mu/\mu \le \omega) = Pr(X \le \omega\sigma + \mu)$$

This is,

$$G(\omega) = \int_{-\infty}^{\omega\sigma+\mu} 1/\sigma\, 2\pi \exp[-(x-\mu)^2/2\sigma^2]dx.$$

If we change the variable integration by writing $y = (x-\mu)/\sigma$, then

$G(\omega) = \int^{\omega}_{-\infty} 1/2\pi\, e^{-y2/2}\, dy.$

Accordingly, the p.d.f. $g(\omega)= G'(\omega)$ of the continuous- type random variable W is

$g(\omega) = 1/2\pi\, e^{-w2/2}$, $-\infty < \omega < \infty$. Thus W is n(0,1),

Theorem 2. If the random variable X is $n(\mu,\sigma^2)$, $\sigma^2>0$, then the random variable $V= (X-\mu)^2/\sigma^2$ is X^2 (1).

Proof. Because $V=W^2$, where $W=(X-\mu)/\sigma$ is n (0,1), the distribution function $G(\upsilon)$ of V is, for $\upsilon \geq 0$,

$G(\upsilon) = Pr(W^2 \leq \upsilon) = Pr(-\sqrt{\upsilon} \leq W \leq \sqrt{\upsilon})$.

That is,

$G(\upsilon) = 2\, {}_0\int^{\sqrt{\upsilon}} 1/2\pi\, e^{-w2/2}\, d\omega, 0 \leq \upsilon,$

And $G(\upsilon)=0$, $\upsilon<0$.

If we change the variable of integration by writing ω= y, then

$G(\upsilon) = {}_0\int^{\upsilon} 1/\sqrt{2\pi y}\, e^{-y/2}\, dy, \quad 0 \leq \upsilon.$

Hence the p. d. f. $g(\upsilon) = G'(\upsilon)$ of the continuous-type random variable V is,

$G(\upsilon) = (1/\sqrt{\pi^2})\, \upsilon^{1/2-1}\, e^{-\upsilon/2}, \; 0 < \upsilon < \infty,$

=0 elsewhere.

Since $g(\upsilon)$ is p. d. f. and hence

${}_0\int^{\infty} g(\upsilon)\, d\upsilon = 1,$

it must be that $\Gamma(1/2) = \sqrt{\pi}$ and thus V is X^2 (1).

4.9 The Bivariate Normal Distribution

Let us investigate the function

$$f(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho_2}}e^{-q/2}, \qquad -\infty<x<\infty, -\infty<y<\infty,$$

Where, with $\sigma_1>0, \sigma_2>0$, and $-1<\rho<1$,

$$q = \frac{1}{1-\rho_2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right) + \left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]$$

At this point we do not know that the constant $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$,and ρ represent parameters of a distribution .As a matter of fact, We do not know that $f(x,y)$ has the properties of a joint p.d.f . It will now be shown that:

(a) $f(x,y)$ is a joint p.d.f

(b) X is $n(\mu_1,\sigma_1^2)$ and Y is $n(\mu_2,\sigma_2^2)$

(c) ρ is the correlation coefficient of X and Y

A joint p.d.f of this form is called a bivariate normal p.d.f.,and the random variables X and Y are said to have a bivariate normal distribution

Example: Let us assume that in a certain population of married couples the height X_1 of a husband and the height x_2 of the wife have a bivariate normal distribution with parameters μ_1=5.8 feet, μ_2=5.3 feet, $\sigma_1=\sigma_2$=0.2 foot, and ρ=0.6.The conditional p.d.f. of X_2, given x_1=6.3, is normal with mean 5.3+(0.)(6.3-5.8)=5.6 and standard deviation(0.2) 1-0.36=0.16.Accordingly, given that the height of the husband is 6.3 feet , the probability that his wife has a height between 5.28 and 5.92 feet is

$\Pr(5.28 < X_2 <5.92/x_1 = 6.3) =N(2) - N(-2) =0.955.$

The moment- generating function of a bivariate normal distribution can be determined as follows. We have

$$M(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{t_1x+t_2y} f(x,y)dx\,dy$$

$$= \int_{-\infty}^{\infty} e^{t_1x} f_1(x)\left[\int_{-\infty}^{\infty} e^{t_2y} f(y/x)\,dy\right] dx$$

for all real values of t_1 and t_2. The integral within the brackets is the moment-generating function of the conditional p.d.f. $f(y/x)$.Since

$f(y/x)$ is a normal p.d.f. with mean $\mu_2 + \rho(\sigma_2/\sigma_1)(x-\mu_1)$ and variance $\sigma_2^2(1-\rho^2)$, then

$$\int_{-\infty}^{\infty} e^{t_2y} f(y/x)\,dy = \exp\{t_2[\mu_2+\rho(\sigma_2/\sigma_1)(x-\mu_1)] = t^2{}_2\sigma^2{}_2(1-\rho^2)/2\}$$

Accordingly ,$M(t_1,t_2)$ can be written in the form

$$\exp\{t_2\mu_2-t_2\rho(\sigma_2/\sigma_1)\mu_1+t^2{}_2\sigma^2{}_2(1-\rho^2)/2\} \int_{-\infty}^{\infty} \exp[(t_1+t_2\rho(\sigma_2/\sigma_1)x\} f_1(x)\,dx.$$

But $E(e^{tx}) = \exp[\mu_1t+(\sigma^2{}_1t^2)/2]$ for all real values of t.Accordingly,if we set $t = t_1+t_2\rho(\sigma_2/\sigma_1)$,we see that $M(t_1,t_2)$ is given by

$$\exp\{t_2\mu_2-t_2\rho(\sigma_2/\sigma_1)\mu_1+t^2{}_2\sigma^2{}_2(1\rho^2)/2)+\mu_1(t_1+t_2\rho(\sigma_2/\sigma_1)+\sigma^2{}_1(t_1+t_2\rho(\sigma_2/\sigma_1))^2/2\}$$

or, equivalently,

$$M(t_1,t_2) = \exp(\mu_1 t_1+\mu_2 t_2+(\sigma^2_1 t^2_1+2\rho\sigma_1\sigma_2 t_1 t_2+\sigma^2_2 t^2_2)/2).$$

It is interesting to note that if, in this moment- generating function $M(t_1,t_2)$,the correlation coefficient ρ is set equal to zero, then

$$M(t_1,t_2) = M(t_1,0)\, M(0,t_2).$$

Thus X and Y are stochastically independent when $\rho=0$, If, conversely, $M(t_1,t_2) \equiv M(t_1,0)M(0,t_2)$, we have $e^{\rho\sigma_1\sigma_2 t_1 t_2}=1$.Since each of σ_1 and σ_2 is positive, then $\rho=0$.

EXERCISES

(1) If $P(C_1)>0$ and if $C_2,C_3,C_4, \ldots$ are mutually disjoint sets, show that $P(C_2 \cup C_3 \cup \ldots/C_1) = P(C_2/C_1)+P(C_3/C_1)+\ldots$.

(2) Prove that

$$P(C_1\cap C_2\cap C_3\cap C_4) = P(C_1)P(C_2/C_1)P(C_3/C_1\cap C_2)P(C_4/C_1\cap C_2\cap C_3).$$

(3) A hand of 13 cards is to be dealt at random and without replacement from an ordinary deck of playing cards. Find the conditional probability that there are at least three kings in the hand relative to the hypothesis that the hand contains at least two kings.

(4) A bowl contains 10 chips. Four of the chips are red, 5 are white, and 1 is blue. If 3 chips are taken at random and without replacement, compute the conditional probability that there is 1 chip of each color relative to the hypothesis that there is exactly 1 red chip among the 3.

(5) Let X_1 and X_2 have the joint p.d.f. $f(x_1,x_2)=x_1+x_2$, $0<x_1<1$, $0<x_2<1$,zero else where . Find the conditional mean and variance of X_2 given $X_1=x_1, 0<x_1<1$.

(6) Let $f(x_1,x_2)=21\, x_1^2 x_2^3$, $0<x_1<x_1<1$, zero else where, be the joint p.d.f . of X_1 and X_2. Find the conditional mean and varience of X_1, given $X_2=x_2, 0<x_2<1$.

(7) If X_1 and X_2 are random variables of the discrete type having p.d.f. $f(x_1,x_2) = (x_1+2x_2) / 18$,$(x_1,x_2 = (1,1),(1,2),(2,1),(2,2)$, zero elsewhere, determine the conditional mean and varience of X_2, given $X_1 = x_1$, $x_1=1$ or 2

(8) Let X_1 and X_2 have the joint p.d.f. $f(x_1,x_2)$ described as follows:

(x_1,x_2)	(0,0)	(0,1)	(1,0)	(1,1)	(2,0)	(2,1)
$f(x_1,x_2)$	1/18	3/18	4/18	3/18	6/18	1/18

and $f(x_1,x_2)$ is equal to zero elsewhere. Find the two marginal probability, density functions and the two conditional means.

(9) Let the random varibales X and Y have the joint p.d.f

(a) $f(x,y) = \frac{1}{3}$, (x,y) = (0,0), (1,1), (2,2), Zero elsewhere

(b) $f(x,y) = \frac{1}{3}$, (x,y) = (0,2), (1,1), (2,0), Zero elsewhere

(c) $f(x,y) = \frac{1}{3}$, (x,y) = (0,0), (1,1), (2,0), Zero elsewhere

In each case compute the correlation coefficient of X and Y

(10) Let X and Y have the joint p.d.f described as follows.

(x , y)	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)
$f(x,y)$	$\frac{2}{15}$	$\frac{4}{15}$	$\frac{3}{15}$	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{4}{15}$

and $f(x,y)$ is equal to zero elsewhere, Find the correlation coefficient ρ

(11) Let $f(x,y) = 2, 0 < x<y, 0 < y < 1$, zero elsewhere, be the joint p.d.f. of X and Y. Show that the conditional means are, respectively, $(1+x)/2$, $0<x<1$, and $y/2$, $0 < y < 1$. Show that the correlation coefficient of X and Y is $\rho = ½$.

(12) Show that the random vaiables X_1 and X_2 with join p.d.f $f(x_1,x_2)=12x_1x_2(1-x_2), 0<x_1<1, 0<x_2<1$, zero elsewhere are stochastically independent.

(13) If the random variables X_1,X_2 have the join p.d.f $f(x_1,X_2)=2e^{-x1-x2}$, $0<x_1,x_2,0<x_2<\infty$, zero elsewhere, show X_1 and X_2 are stochastically dependent.

(14) Find Pr $(0<X_1<1/3, 0<X_2<1/3)$ if the random variable X_1 and X_2 have the joint p.d.f $f(x_1,x_2) = 4x_1(1-x_2), 0<x_1<1, 0<x_2<1$, zero elsewhere.

(15) If the moment -generating function of a random variable X is $(1/3+2/3^{et})^5$, find Pr(X = 2or3)

(16) The moment generating function of a random variable X is (2/3+1/3et)9. Show that

$$\Pr(\mu-2\sigma<X<\mu+2\sigma) = {}_{x=1}{}^{5}\Sigma(9/x)(1/3)x(2/3)^{9-x}$$

(17) If X is b(n,p), show that

$E(X/n)=p$ and $E[(X/n-p)^2] = p(1-p)/n$

(18) Let Y be the number of success in n independent repetitions of a random experiment having the probability of success p=2/3. If n=3, computer Pr (2<Y); if n=5, compute Pr(3≤Y)

(19) Let X be b(2,p) and let Y be b(4,p). If Pr(X≥1)5/9, find Pr (Y≥1).

(20) Show that the moment generation function of the negative binomial distribution is $M(t) = pr[1-(1-p)et]^{-r}$,. Find the mean and variance of this distribution. Hint. In the summation respesenting M(t) , make use of the MacLaurin,s series for $(1-\omega)^{-r}$

(21) If a fair coin is tossed at random five independent times, find the conditional probability of five heads relative to the hypothesis that there are at least four heads.

(22) If the random variable X has a poisson distribution such that

Pr (X =1) = Pr (X=2), find Pr(X=4).

(23) The moment generating function of a random variable X is $e^{4(et-1)}$. Show that $Pr(\mu - 2\sigma < X < \mu + 2\sigma) = 0.931$

(24) Compute the measures of skewness and kurtosis of the poisson distribution with mean μ

(25) Let X and Y have the joint p.d.f. $f(x,y) = e^{-2} / [x!\,(y-x)!]$

y = 0, 1,2,.....; x = 0.1 ...... y, Zero elsewhere

(a) Find the moment-generating function $M(t_1,t_2)$ of this joint distribution.

(b) Compute the means, the variances, and the correlation coefficient of X and Y.

(c) Determine the conditional mean E(X ≥ y). Hint,

(26) If $(1-2t)^{-6}$, t<1/2 is the moment -generating function of the random variable X, find Pr (X<5.23).

(27). If X is χ2(5), determine the constants c and d so that Pr(C<X<d) = 0.95 and Pr(X<c)=0.025

(28) Let X have a gamma distribution with p.d.f f(x)=1/β2 xe-x/β,0<x<∞, zero elsewhere. If x=2 is the unique mode of the distribution, find the parameter β and Pr (X<9.49).

(29). Compute the measures of skew ness and kurtosis of a gamma distribution with parameters α and β.

(30) Let X have the uniform distribution with p.d.f. $f(x)=0,1<x<1$, zero elsewhere. Find the distribution function of Y=-2in X. What is the p.d.f of Y?

31. If $N(x) = {}_{-\infty}\int^{x} 1/\sqrt{2\pi}\, e^{-w2/2}\, d\omega$,

show that N(-x)=1-N(x).

32. If X is n (75, 100), find Pr (X-60) and Pr (70<X<100).

33. If X is n (μ, σ^2), find b so that Pr $[-b<(X-\mu)/\sigma<b] = 0.90$.

34. If X is n (μ, σ^2), show that E $(|X-\mu|)=\sigma^2/\pi$.

35. Let the random variable X have the p. d. f.

$f(x)=2/2\pi\, e^{-x2/2}$, $0<x<\infty$, zero elsewhere.

Find the mean and variance of X. Hint. Compute E (X) directly and E (X^2) by comparing that integral with the integral representing the variance of a variable that is n (0, 1).

36. Let X be n (5, 10). Find Pr $[0.04<(X-5)^2< 38.4]$.

37. Let X and Y have a bivariate normal distribution with parameters

$\mu_1 = 3$, $\mu_2 = 1$, $\sigma^2_1 = 16$, $\sigma_2^2 = 25$, and $\rho = 3/5$. Determine the following probabilities :

(a) Pr(3<Y<8).

(b) Pr(3< Y <8/x=7).

(c) Pr(-3< X <3).

(d) Pr(-3< X<3/y=-4).

38. Let X and Y have a bivariate normal distribution with parameters

$\mu_1 = 5$, $\mu_2 = 10$, $\sigma^2_1 = 1$, $\sigma^2_2 = 25$, and $\rho>0$.

If Pr(4 < Y<16/x=5)=0.954, determine ρ.

UNIT V

DISTRIBUTIONS OF FUNCTIONS OF RANDOM VARIABLES

TABLE OF CONTENTS

5. 1 SAMPLING THEORY

Definition :

A function of one or more random variables that does not depend upon any *unknown* parameter is called a statistic.

Definition :

Let $X_1,X_2...,X_n$ denote n mutually stochastically independent random variables, each of which has the same but possibly unknown p.d.f. f(x); that is, the probability density functions of $X_1,X_2,....,X_n$ are , respectively, $f_1(x_1) = f(x_1),f_2(x_2) = f(x_2), ... ,f_n(x_n) = f(x_n)$, so that the joint p.d.f. is $f(x_1)f(x_2)...f(x_n)$. The random variables $X_1,X_2,...,X_n$ are then said to constitute a random sample from a distribution that has p.d.f. f(x).

Definition :

Let X_1 ,X_2,....,Xn denote a random sample of size n from a given distribution . The statistic

$$\bar{X} = \frac{X_1+X_2+\ldots\ldots+X_n}{n} = \sum_{i=1}^{n} \frac{X_i}{n}$$

is called the mean of the random sample, and the statistic

$$S^2 = \sum_{i=1}^{n} \frac{(X_i-\bar{X})^2}{n} = \sum_{i=1}^{n} \frac{X_i^2}{n} - \bar{X}^2$$

is called the variance of the random sample.

Example :

Let the random variable Y be distributed uniformly over the unit interval $0<y<1$; that is the distribution function of Y is

$$G(y) = 0, y\leq 0$$
$$= y, 0<y<1,$$
$$= 1, 1\leq y$$

Suppose that F(x) is a distribution function of the continuous type which is strictly increasing when $0 < F(x)<1$. If we define the random variable X by the relationship $Y = F(X)$, we now show that X has a distribution which corresponds to F(x). If $0<F(x)<1$, the inequalities $X\leq x$ and $F(X) \leq F(x)$ are equivalent. Thus , with $0 <F(X)<1$, the distribution function of X is

$$Pr(X\leq x)= Pr(F(x)\leq F(x)]=Pr[Y\leq F(x)]$$

because $Y =F(X)$. However ,$Pr(Y\leq y)= G(y)$, so we have

$$Pr(X\leq x)= G[F(x)]=F(x).\ 0<F(x)<1$$

That is the distribution function of X i F(x).

This result permits us to simulate random variables of different types.

5.2 TRANSFORMATIONS OF VARIABLES OF THE DISCRETE TYPE

An alternative method of finding the distribution of a function of one or more random variables is called the change of variable technique.

Let X have the poisson p.d.f

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}, \qquad x = 0,1,2,\ldots\ldots$$

= 0 elsewhere.

Let A denote the space A = {x;x = 0,1,2,3.....}, so that A is the set where f(x)>0. Define a new random variable Y by Y = 4X. We wish to find the p.d.f. of Y by the change-of-variable technique. Let y = 4x. We call y = 4x a transformation from x to y, and we say that the transformation maps the space A on to the space B ={y;y=0,4,8,12.....}. The space B is obtained by transforming each point in a in accordance with y = 4x.

The p.d.f . g(y) of the discrete type

$$g(y) = Pr(Y=y) = Pr(X=y/4) = \frac{\mu^{y/4}e^{-\mu}}{(y/4)!}, \quad y = 0,4,8.....$$

0 = elsewhere.

Example . Let X have the binomial p.d.f.

$$f(x) = \frac{3!}{x!(3-x)!}\left(\frac{2}{3}\right)^x \left(\frac{1}{3}\right)^{3-x}, \quad x = 0,1,2,3,$$

= 0 elsewhere.

We seek the p.d.f. g(y) of the random variable $Y=X^2$. The transformation $y = u(x)=x^2$ maps A = {x;x=0,1,2,3} on to B={y;y=0,1,4,9}. In general, $y=x^2$ does not define a one-to-one transformation; here, how ever, it does, for there are no negative values of x in A={x; x=0,1,2,3}. That is, we have the single -valued in verse function $x = w(y) = \sqrt{y}$ (not $-\sqrt{y}$), and so

$$g(y) = f(\sqrt{y}) = \frac{3!}{(\sqrt{y})!\,(3-\sqrt{y})!}\left(\frac{2}{3}\right)^{\sqrt{y}}\left(\frac{1}{3}\right)^{3-\sqrt{y}}, \quad y = 0,1,4,9,$$

= 0 elsewhere.

Example

Let X_1 and X_2 be two stochastically independent random variables that have Poisson distributions with means μ_1 and μ_2, respectively.

The joint p.d.f of X_1 and X_2 is

$$\frac{\mu_1^{x_1}\mu_2^{x_2}e^{-\mu_1-\mu_2}}{x_1!x_2!} \quad x_1=0,1,2,3,...., \quad x_2=0,1,2,3,..,$$

and is zero elsewhere. Thus the space A is the set of points (x_1,x_2), where each of x_1 and x_2 is a nonnegative integer. We wish to find the p.d.f of $Y_1=X_1+X_2$. If we use the change of variable technique , we need to define a second random variable Y_2. Let us choose Y_2 in such a way that a simple

one-to-one transformation .For example, take $Y_2=X_2$. Then $y_1=x_1+x_2$ and $y_2=x_2$ represent one-to-one transformation that maps A on to

$$B=\{(y_1,y_2); y_2=0,1,\ldots,y_1 \text{ and } y_1=0,1,2,\ldots\}.$$

Note that, if$(y_1,y_2)\in B$, then $0< y_2 < y_1$.The inverse functions are given by $x_1=y_1-y_2$ and $x_2=y_2$. Thus the joint p.d.f. of Y_1 and Y_2 is

$$g(y_1,y_2) = \frac{\mu_1^{y_1-y_2}\mu_2^{y_2}e^{-\mu_1-\mu_2}}{(y_1-y_2)!y_2!} \qquad (y_1,y_2)\in B,$$

and is zero elsewhere . Consequently , the marginal p.d.f of Y_1 is given by

$$g_1(y_1) = \sum_{y_2=0}^{y_1} g(y_1,y_2)$$

$$= \frac{e^{-\mu_1-\mu_2}}{y_1!} \sum_{y_2=0}^{y_1} \frac{y_1!}{(y_1-y_2)!y_2!} \mu_1^{y_1-y_2}\mu_2^{y_2}$$

$$= \frac{(\mu_1+\mu_2)^{y_1}e^{-\mu_1-\mu_2}}{y_1!} \qquad y_1=0,1,2,\ldots,$$

and is zero elsewhere. That is , $Y_1=X_1+X_2$ has a Poisson distribution with parameter $\mu_1+\mu_2$.

5. 3 TRANSFORMATIONS OF VARIABLES OF THE CONTINOUS TYPE

Example.

Let X be the random variable of the continuous type, having p.d.f.

$f(x)=2x$, $0<x<1$, $f(x)=0$ elsewhere

Here A is the space $\{x;0<x<1\}$where $f(x)>0$. Define the random variable y by $y=8x^3$, and consider the transformation $y=8x^2$. Under the transformation $y=8x^2$, the set A is mapped on to the set $B=\{y;0<y<8\}$, and moreover the transformation is 1 to 1. for every $0<a<b<8$, the event $a<Y<b$ will occur when and only when the event $\frac{1}{2}\sqrt[3]{a}<x<1/2\sqrt[3]{b}$ occurs because there is a one to one correspondence between the points of a and b.

Thus

$Pr(a<y<b) = Pr(1/2\ \sqrt[3]{a} < X < \frac{1}{2}\sqrt[3]{b})$

$$\int_{\sqrt[3]{a/2}}^{\sqrt[3]{b/2}} 2x\,dx$$

By changing the variable of integration from x to y by writing $y = 8x^3$ or $x=1/2\ \sqrt[3]{y}$.

$$\frac{dx}{dy} = \frac{1}{6y^{2/3}}$$

and accordingly, we have

$$Pr(a<Y<b) = {}_a\int^b 2\frac{(\sqrt[3]{y})}{2}\left(\frac{1}{6y^{2/3}}\right)dy$$

$$= {}_a\int^b \frac{1}{6y^{1/3}}\,dy$$

Since this is true for every $0<a<b<8$, the p.d.f g(y) of Y is the integrand; that is,

$$g(y) = \frac{1}{6y^{1/3}} \quad 0<y<8,$$

$= 0$ elsewhere

$g(y) = 0$ otherwise

Example: Let X have the p.d.f.

$f(x) = 1,\ 0<x<1,$

$= 0$ elsewhere.

We have to show that the random variable $Y= -2 \ln X$ has a chi-square distribution with 2 degrees of freedom. Here the transformation $y=\mu(x)= -2 \ln x$ so that $\omega(y) = e^{-y/2}$. The space **A** is $A = \{x;0<x<1\}$ which the one to one transformation $y = -2 \ln x$ maps onto **B**$=\{y;0<y<\infty\}$. The Jacobian of the transformation is

$$J = \frac{dx}{dy} = \omega^1(y) = \frac{1}{2}e^{-y/2}$$

According, the p.d.f g(y) of $Y = -2 \ln X$ is

$g(y) = f(e^{-y/x})|J| = 1/2\ e^{-y/2},\ 0<Y<\infty,$

= 0 elsewhere, a p.d.f that is chi-square with 2 degrees of freedom. This method of finding the p.d.f of a function of one random variable of the continuous type will now be extended to function of two random variables of this type. Again, only functions that define a one-to-one transformation will be considered at this time. Let $y_1=\mu_1(x_1,x_2)$ and $y_2=\mu_2(x_1,x_2)$ define a one-to-one transformation that maps a (two-dimensional) set A in the x_1,x_2 plane onto a (two - dimensional) set A in the y_1,y_2- plane. If we express each of x_1 and x_2 in terms of y_1 and y_2, we can write $x_1=\omega_1(y_1,y_2)$, $x_2=\omega_2(y_1,y_2)$. The determinant of order 2,

$$\begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$$

is called the jacobian of the transformation and will be denoted by the symbol J

Example

Let the random variable X have the p.d.f

$f(x)=1,\ 0<x<1,$

= 0 elsewhere

and let X_1,X_2 denoted a random sample from this distribution. The joint p.d.f of X_1 and X_2 is then

$\varphi(x_1,x_2)=f(x_1)(x_2)=1,\ 0<x_1<1<,0,x_2<1,$

= 0 elsewhere

Consider the two random variables $Y_1=X_1+X_2$ and $Y_2=X_1-X_2$, we wish to find the join p.d.f of Y_1 and Y_2. Here the two dimensional space A in the x_1,x_2 plane is that of Example 3 of this section. The one-two-one transformation $y_1=x_1+x_2$, $y_2=x_1-x_2$ maps **A** onto the space **B** of that example. Moreover, the Jacobian of that transformation has been shown to be J = -1/2. Thus

$$g(y_1,y_2) = \varphi[1/2(y_1+y_2),1/2\ (y_1-y_2)]|J|$$

$$= f[1/2(y_1+y_2)]f[1/2(y_1-y_2)|J|=1/2 \qquad (y_1,y{=}2{=}\mathbf{B})$$

$$= 0 \text{ elsewhere}$$

Because **B** is not a product space, the random variable Y_1 and Y_2 are stochastically dependent.

The Marginal p.d.f of Y_1 is given below

$g_1(y_1) = {}_{-y_1}\int^{y_2} 1/2\, dy_2 = y_1 \leq 1$

$= {}_{y2-2}\int^{2-y_1} ½\, dy_2 = 2-y,\ 1\leq y_1 \leq 2,$

$= 0$ elsewhere

In a similar manner , the marginal p.d.f $g_2(y_2)$ is given by

$g_2(y_2) = {}_{-y2}\int^{y1} 1/2\, dy_2 = y1,\ 0<y1\leq 1$

$= {}_{(y1-2)}\int^{2-y2} ½\, dy_1 = 1-y_2,\ 1<y_1<2,$

$= 0$ elsewhere

Example

Let $y_1=(x_1-x_2)$ where x_1 and x_2 are stochastically independent random variables, each being $X_2(2)$. The join p.d.f of X_1 and X_2 is

$f(x_1)(x_2) = 1/4 \exp(-x_1\_x_2),\ 0<x1<\infty,\ 0<x2<\infty.$

$= 0$ elsewhere

Let $X_1=X_2$ so that $y1=1/2(x_1-x_2), y_2=x_2$, or $x_1=2y_1+y_2$, $x_2=y_2$ define a one-to-one transformation from $\mathbf{A}=\{(x_1,x_2);\ 0<x1<\infty,\ 0<x2<\infty\}$ onto $\mathbf{B}=\{(y_1,y_2);-2y_1<y_2$ and $0<y_2,\ -\infty<y_2<\infty)\}$, The jacobian of the transformation is

$$J = \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} = 2$$

Hence the join p.d.f of Y_1 and Y_2 is

$g(y_1,y_2) = 2\ {}_{e}{-y_1-y_2} \qquad (y_1,y_2)\in B$

$= 0$ elsewhere

Thus the p.d.f of Y_1 is given by

$g_1(y_1) = {}_{-2y_1}\int^{\infty} ½\ {}_{e}{-y1-y_2}\, dy_2 = ½\, ex_1, \qquad -\infty<y1<0$

$= {}_0\int^{\infty} 1/2\ {}_{e}{-y_1-y_2}\, dy_2 = 1/2\ {}_{e}{-y_2}, \qquad 0\leq y1<\infty,$

or

$g_1(y_1) = 1/2\, e^{-y_1}, \qquad -\infty<y_1<\infty$

This p.d.f is called the double exponential p.d.f

5. 4 THE t AND F DISTRIBUTIONS

Let W denote a random variable that is n(0,1); Let V denote a random variable that is $\chi^2(\gamma)$; and let W and V be stochastically independent. Then the joint p.d.f of W and that of V or

$\varphi(\omega,v) = 1/\sqrt{2\pi}\ e^{-w2/2}\ 1/r(r/2)2r/2\ vr/2-1\ e^{-r/2}$, $\quad -\infty<\omega<\infty,\ 0<v<\infty,$

= 0 elsewhere

Define a new random variable T by writing

$T=W/\sqrt{V/r}$

The change-of-variable technique will be used to obtain the p.d.f $g_1(t)$ of Y. The equations.

$t=\omega/\sqrt{v/r}$ and $u=v$

define a one-to-one transformation that maps A $=\{(W,v);-\infty<\infty\}$. Since $\omega = t\sqrt{u}/\sqrt{r},\ v=u$, the absoulte value of the Jacobian of the transformation is $|J| = \sqrt{u}/\sqrt{r}$. Accordingly the joint p.d.f of T and U = V is given by

$g(t,u) = \varphi(t\sqrt{u/r},\ u)\ |J|$

$$= \frac{u^{r/2-1}}{\sqrt{2\pi}\ \sqrt{r/2}\ 2^{r/2}} \exp\{-u\ \frac{(1/t^2/r)}{2}\} \frac{\sqrt{u}}{\sqrt{r}}$$

$= 1/\sqrt{2\pi}(r/2)2^{r/2}\ ur/2-1\ \exp[-u/2\ (1+t2/r)]$

= 0 elsewhere $\quad -\infty<t<\infty,\ 0<u<\infty$

The marginal p.d.f of T is then

$$g1(t) = {}_{-\infty}\int^{\infty} g(t,u)\ du$$

$$= {}_0\int^{\infty} \frac{u(r+1)\ 2-1 \quad \exp\ -y/2\ (1+t^2/r)\ dx}{2\sqrt{\pi r}\ \ \sqrt{r/2}\ \ 2^{r/2}}$$

In this integral let

$x = u\,[1+t^2/r]/2$ and then

$$g_1(t) = {}_0\int^{\infty} \frac{2z/1+t^2/r)\quad e^{-x}\quad 2}{\sqrt{2\pi r}\quad \sqrt{r/2}\ 2^{r/2}\quad \overline{1+t/r}\ dx} \qquad -\infty,<t<\infty$$

Thus if W is n(0,1) is n(0,1), if V is $\chi^2(r)$, and if W and V are stochastical independent, then

$T = W/\sqrt{V/r}$

5.5 Extensions of the change-of-variable Technique

Consider an integral of the form

$\int \ldots_A \int \varphi(x_1, x_2, \ldots, x_n)\, dx_1\, dx_2 \ldots dx_n$

taken over a subset A of an n- dimensional space **A**. Let

$y_1 = u_1(x_1, x_2, \ldots, x_n),\ y_2 = u_2(x_1, x_2, \ldots, x_n), \ldots$

$y_n = u_n(x_1, \ldots, x_n),$

together with the inverse functions.

$x_1 = w_1(y_1, y_2, \ldots, y_n),\ x_2 = w_2(y_1, y_2, \ldots, y_n) \ldots$

$x_n = w_n(y_1, y_2, \ldots, y_n)$

Define a one to one transformation that maps **A** and **B** in the $y_1, y_2 \ldots y_n$ space (and hence maps the subset A of **A** and on to a subset **B** of **B**). Let the first partial derivatives of the inverse functions be continuous and let the n by n determinant (called the Jacobian)

$$F = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_n} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & \cdots & \frac{\partial x_2}{\partial y_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial x_n}{y_1} & \frac{\partial x_n}{\partial y_2} & & \frac{\partial x_n}{\partial y_n} \end{vmatrix}$$

not vanish identically in **B** . Then

$\int_A \ldots \int \varphi(x_1, x_2, \ldots, x_n) dx_1, dx_2 \ldots dx_n$

$= \int_B \ldots \int \varphi[w_1(y_1, \ldots, y_n), w_2(y_1 \ldots, y_n) \ldots w_n(y_1 \ldots, y_n)] \times |J|\, dy_1\, dy_2 \ldots dy_n$

The joint p.d.f. of the random variables $Y_1 = u_1(X_1, X_2, \ldots X_n)$,

$Y_2 = u_2(X_1,X_2,...X_n),....,Y_n = u_n(X_1,X_2....X_n)$- where the joint p.d.f of $X_1,X_2,...X_n$ is $\varphi(x_1,....,x_n)$ is given by

$$g(y_1,y_2,...y_n) = |J|\ \varphi[w_1(y_1,....,y_2),...,w_n(y_1,...,y_n)],$$

when $(y_1,y_2,...,y_n) \in \mathbf{B}$, and is zero elsewhere.

Example 1:

Let $X_1, X_2,..,X_{k+1}$ be mutually stochastically independent random variables, each having a gamma distribution with $\beta=1$. The joint p.d.f of these variables may be written as

$$\varphi(x_1,x_2,...,x_{k+1}) = \prod_{i=1}^{k+1} 1/\Gamma(\alpha_i) x_i^{\alpha_i-1} e^{-x_i},\ 0<x_i<\infty,$$
$$= 0 \text{ else where}$$

Let

$$Y_i = \frac{X_i}{X_1+X_2+...+X_{k+1}},\quad i=1,2,...,k,$$

and $Y_{k+1}=X_1+X_2+...+X_{k+1}$ denote k+1 new random variables The associated transformation maps $A=\{(x_1,...,x_{k+1});\ 0<x_i<\infty,\ i=1,...,k+1\}$ onto the space.

$B=\{(y_1,...,y_k,y_{k+1});0<y_i,i=1,...,k,$

$y_1+...+y_k<1, 0<y_{k+1}<\infty\}.$

The single-valued inverse functions are $x_1 = y_1y_{k+1},...,x_k = y_ky_{k+1}$,

$x_{k+1} = y_{k+1}(1-y_1-...-y_k)$, so that the Jacobian is

$$J = \begin{vmatrix} y_{k+1} & 0 & \cdots & 0 & y_1 \\ 0 & y_{k+1} & \cdots & 0 & y_2 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & y_{k+1} & y_k \\ -y_{k+1} & -y_{k+1} & \cdots & -y_{k+1} & (1-y_1-...-y_k) \end{vmatrix} = Y^k_{k+1}$$

Hence the joint p. d. f. of $Y_1, ..., Y_k, Y_{k+1}$ is given by

$$\frac{y_{k+1}^{\alpha_1+\dots+\alpha_{k+1}-1}\, y_1^{\alpha_1-1} \cdots y_k^{\alpha_k-1} (1-y_1-\dots-y_k)^{\alpha_{k+1}-1} e^{-y_{k+1}}}{\Gamma(\alpha_1)\dots\Gamma(\alpha_k)\,\Gamma(\alpha_{k+1})}$$

provided that $(y_1, \dots, y_k, y_{k+1}) \in$ **B** and is equal to zero elsewhere.

The joint p. d. f. of $Y_1, \dots, Y_k$ is

$$g(y_1, \dots, y_k) = \frac{\Gamma(\alpha_1+\dots+\alpha_{k+1})\, y_1^{\alpha_1-1} \dots y_k^{\alpha_k-1} (1-y_1-\dots-y_k)^{\alpha_{k+1}-1}}{\Gamma(\alpha_1 \dots \Gamma(\alpha_{k+1})}$$

When $0<y_i,\ i=1, \dots, k,\ y_1+\dots+y_k<1$, while the function g is equal to zero elsewhere. Random variables $Y_1, \dots, Y_k$ that have a joint p.d.f. of this form are said to have a Dirichlet distribution with parameters $\alpha_1, \dots, \alpha_k, \alpha_{k+1}$, and any such $g(y_1,\dots,y_k)$ is called a Dirichlet p.d.f. It is seen, in the special case of k=1, that the Dirichlet p.d.f. becomes a beta p.d.f. Moreover, it is also clear from the joint p.d.f. of $Y_1,\dots,Y_k, Y_{k+1}$ that Y_{k+1} has a gamma distribution with parameters $\alpha_1+\dots+\alpha_k+\alpha_{k+1}$ and $\beta=1$ and that Y_{k+1} is stochastically independent of $Y_1, Y_2, \dots, Y_k$.

Now, let X have the Cauchy p.d.f.

$$f(x) = 1/\pi(1+x^2), \quad -\infty<x<\infty,$$

and let $Y=X^2$. We seek the p.d.f. g(y) of Y. Consider the transformation $y=x^2$. This transformation maps the space of X, $A=\{x;\ -\infty<x<\infty\}$, onto $B=\{y;0\le y<\infty\}$. However, the transformation is not one-to-one. To each $y\in B$, with the exception of y=0, there correspond two points $x\in A$. for example, if y=4, we may have either x = 2 or x = -2. In such an instance, we represent A as the union of two disjoint sets A_1 and A_2 such that $y=x^2$ defines a one-to-one transformation that maps each of A_1 and A_2 onto B. If we take A_1 to be $\{x;-\infty<x<0\}$ and A_2 to be $\{x;0\le x<\infty\}$, we see that A_1 is mapped onto $\{y;0<y<\infty\}$, where as A_2 is mapped onto $\{y;0\le y<\infty\}$, and these sets are not the same.

Take $A1=\{x;-\infty<x<0\}$ and $A2=\{x;0<x<\infty\}$. Thus $y = x^2$, with the inverse $x = -\sqrt{y}$, maps A_1 onto $B = \{y;0<y<\infty\}$ and the transformation is one-to-one. Moreover, the transformation y =

x^2, with inverse $x = -\sqrt{y}$, maps A_2 onto **B** $= \{y; 0<y<\infty\}$ and the transformation is one-to-one. Consider the probability $Pr(Y \in B)$, where $B \subset \mathbf{B}$. Let $A_2 = \{x;\ x = -\sqrt{y},\ y \in B\} \subset A_1$ and let $A_4 = \{x\ ;\ x = \sqrt{y},\ y \in A_4$. Thus we have

$$Pr = (Y \in B) = Pr(X \in A_3) + Pr(X \in A_4)$$

$$= \int_{A3} f(x) + \int_{A4} f(x)dx.$$

In the first of these integrals, let $x = -\sqrt{y}$. Thus the Jacobian, say J is $-1/\ 2\sqrt{y}$ moreover, the set A_4 is mapped onto B. In the second integral let $x = \sqrt{y}$. Thus the Jacobian, say J_2, is $1/2\sqrt{y}$; moreover, the set A_4 is also mapped onto B. Finally,

$$Pr(Y \in B) = \int_B f(-\sqrt{y}) \mid 1/2\sqrt{y} \mid dy + \int f(f\sqrt{y}) 1/2\sqrt{y}\, dy$$

$$= \int_B [f(-\sqrt{y})\} f(/\sqrt{y})] 1/2\sqrt{y}\, dy.$$

Hence the p.d.f. of Y is given by

$$g(y) = 1/2\sqrt{y}[f(-\sqrt{y}) + f(\sqrt{y})], \quad y \in B.$$

With f(x) the Cauchy p.d.f. we have

$$g(y) = 1/\pi(1+y)\sqrt{y}, \quad 0<y<\infty,$$

$$= 0 \text{ elsewhere.}$$

Let $\varphi(x_1, x_2, \ldots, x_n)$, be the joint p.d.f. of $X_1, X_2, \ldots, X_n$, which are random variables of the continuous type. Let **A** be the n-dimensional space where $\varphi(x_1,x_2,\ldots,x_n), \ldots, y_n = \mu_n(x_1,x_2,\ldots,x_n)$, which maps **A** onto **B** on the $y_1,y_2,\ldots,y_n$ space. To each point of **A** there will correspond, of course, but one point in **B**; but to a point in **B** there may correspond more than one point in **A**. That is, the transformation may not be one-to-one. Suppose, however, that we can represent **A** as the union of a finite number, say k, of mutually disjoint sets $A_1, A_2, \ldots A_k$ so that.

$$Y_1 = \mu_1(x_1,x_2,\ldots,xn, \ldots, \quad yn = \mu n(x_1,x_2,\ldots,x_n)$$

Define a one-to-one transformation of each A_i onto **B**. Thus, to each point in **B** there will correspond exactly one point in each of $A_1, A_2, \ldots A_k$.

Let $x_1 = \omega_{1i}(y_1,y_2,\ldots,y_n)$,

$x_2 = \omega_{2i}(y_1,y_2,\ldots,yn)$, $i = 1, 2, \ldots, k$,

$\vdots$

$x_n = \omega_{ni}(y_1,y_2,\ldots,y_n)$, denote the k groups of n inverse functions, one group for each of these k transformations. Let the first partial derivatives be continuous and let each

$$J_i = \begin{vmatrix} \frac{\partial \omega_{1i}}{\partial y_1} & \frac{\partial \omega_{1i}}{\partial y_2} & \cdots & \frac{\partial \omega_{1i}}{\partial y_n} \\ \frac{\partial \omega_{2i}}{\partial y_1} & \frac{\partial \omega_{2i}}{\partial y_2} & \cdots & \frac{\partial \omega_{2i}}{\partial y_n} \\ \frac{\partial \omega_{ni}}{\partial y_1} & \frac{\partial \omega_{ni}}{\partial y_2} & \cdots & \frac{\partial \omega_{ni}}{\partial y_n} \end{vmatrix} \quad i = 1, 2, \ldots, k$$

not identically equal to zero in **B** . From a consideration of the probability of the union of k mutually exclusive events and by applying the change of variable technique to the probability of each of these events, it can be seen that the joint p.d.f. of $Y_1 = u_1(X_1, X_2, \ldots, X_n)$, $Y_2 = u_2(X_1, X_2, \ldots, X_n), \ldots$ $Y_n = u_n(X_1, X_2, \ldots, X_n)$ is given by

$g(y_1, y_2, \ldots, y_n) = \sum_{i=1}^{k} |J_i| \, \varphi[w_{1i}(y_1, \ldots, y_n) \ldots w_{ni}(y_1, \ldots, y_n)]$, provided that $(y_1, y_2, \ldots, y_n) \in$ **B** and equals zero elsewhere. The p.d.f. of any Y_i, say Y_1, is then

$$g_1(y_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(y_1, y_2, \ldots, y_n) dy_2 \ldots dy_n.$$

Example 2:

To illustrate the result just obtained , take n =2 and let X_1 , X_2 denote a random sample of size 2 from a distribution that is n(0,1), The joint p.d.f. of X_1 and X_2 is

$$f(x_1, x_2) = 1/2\pi \exp[(-x_1^2 + x_2^2)/2], \; -\infty < x_1 < \infty, -\infty < x_2 < \infty.$$

Let Y_1 denote the mean and let Y_2 denote twice the variance of the random sample. The associated transformation is

$$y_1 = x_1 + x_2/2,$$

$$y_2 = (x_1 - x_2)^2/2$$

The transformation maps **A**$=\{(x_1, x_2); -\infty < x_1 < \infty, -\infty < x_2 < \infty\}$ onto **B** $=\{(y_1 - y_2); -\infty < y_1 < \infty, -\infty < y_2 < \infty\}$. But the transformation is not one-to-one because, to each point in **B** , exclusive of points when $y_2=0$, there correspond two points in **A**. In fact the two groups of inverse functions are

$x_1 = y_1-\sqrt{y_2/2}$ $\quad x_2 = y_1+\sqrt{y_2/2},$

and

$X_1 = y_1+\sqrt{y_2/2}$ $\quad x_2 = y_1-\sqrt{y_2/2}.$

Moreover the set **A** cannot be represented as the union of disjoint sets each of which under our transformation maps onto **B**. Our difficulty is caused by those points of A that lie on the line whose equation is $x_2 = x_1$. At each of these points we have $y_2 = 0$. However, we can define $f(x_1,x_2)$ to be zero at each point where $x_1 = x_2$. We can do this without altering the distribution of probability, because the probability measure of this set is zero. Thus we have a new $A=\{(x_1,x_2); -\infty<x_1<\infty, -\infty<x_2<\infty\}\ x_1 \neq x_2\}$. This space is the union of the two disjoint sets $A_1=\{(x_1,x_2); x_2>x_1\}$ and $A_2=\{(x_2<x1); x_2<x_1\}$./ Moreover our transformation now defines a one–to-one transformation of each Ai,i=1,2, onto the new **B** $=\{(y_1,y_2); -\infty<y_1<\infty, -\infty<y_2<\infty\}$. We can now find the joint p.d.f. say $g(y_1,y_2)$, of the mean Y_1 and twice the variance Y_2 of our random sample.

$$|J_1| = |J_2| = 1/\sqrt{2y_2}. \text{ Thus}$$

$$g(y_1,y_2) = 1/2\pi \exp[-(y_1-\sqrt{y_2/2})^2 - (y_1+\sqrt{y_2/2})2]\ 1/\sqrt{2y_2}$$

$$+ ½\ \pi \exp[-y_1-\sqrt{y_2/2})2/2-(y_1-\sqrt{y_2/2})^2]\ 1/\sqrt{2y_2}$$

$$= \sqrt{2/2\pi}\, e^{-ui2}\ \frac{1}{\sqrt{2\Gamma(1/2)}}\ y_2^{\,½-1}\ e^{-y2/2},\ -\infty<y_1<\infty, -\infty<y_2<\infty\}$$

The mean Y_1 of our random sample is $n(0,1/2)$; Y_2, which is twice the variance of our sample , is $,x^2(1)$; and the two are stochastically independent. Thus the mean and the variance of our sample are stochastically independent.

5.6 The Moment- Generating- function Technique

Let $\varphi(x_1,x_2,x_3.......x_n)$ denote the join p.d.f of the n random variables $X_1, X_2.....X_n$. These random variables any or may not be the items of a random sample from some distribution that has a given p.d.f f(x)_. Let $Y_1=u_1(X_1,X_2,X_3.......X_n)$. We seek $g(y_1)$, the p.d.f of the random variable Y_1. Consider the moment-generation function of Y1. if it exists, it is given by

$M(t)=E(e^{tY_1})=\int_{-\infty}^{\infty} e^{ty_1} g(y_1)dy_1$ in the continuous case.

Example 1

Let the stochastically independent random variables X_1 and X_2 have the same p.d.f

$f(x)=x/6$, $x = 1,2,3$

$= 0$ elsewhere

that is the p.d.f of X_1 is $f(x_1)$ and that of X_2 is $f(x_2)$; and so the joint p.d.f of X_1 and X_2 is

$f(x_1)f(x_2) = x_1,x_2/36 \quad x_1=1,2,3, \; x_2 = 1,2,3$

$= 0$ elsewhere

A probability, such as $Pr(X_1=2, X_2=3)$ can be seen immediately to be $(2)(3)/36=1/6$. Consider a probability such as $Pr(X_1+X_2=3)$. the computation can be made by first observing that the event $X_1+X_2 =3$ is the union exclusive of the events with probability zero of the non mutually exclusive events ($X_1=1$, $X_2=2$) and ($X_1=2$, X_1). The

$Pr(X_1+X_2=3) = Pr(X_1=1,X_2=2)+ Pr(X_1=2,X_2=1)$

$= (1)(2)/36+ (2)(1)/36 = 4/36.$

More generally , let y represent any of the number 2,3,4,5,6. The probability of each of the events $X_1+X_2 = y$,$y = 2,3,4,5,6$ can be computed. Let $g(y) = Pr(X_1+X_2=y)$. Then the table

y	2	3	4	5	6
g(y)	1/36	4/36	10/36	12/36	9/36

gives the values of g(y) for y 2,3,4,5,6, . For all the values of y, g(y), = 0 . Now, define a new random variable Y by $Y = X_1+X_2$, and the we have to calculate the p.d.f .g(y) of this random variable Y. We shall now solve the same problem and by the moment generating function technique.

Now the moment generating function of Y is

$M(t) \quad = \quad E(e^{t(x1+x2)})$

$= \quad E(e^{tX1} . e^{tX2})$

$= \quad E(e^{tX1})Ee(^{tx2})$, Since X_1 and X_2 are stochastically independent.

Theoram 1:

Let $X_1, X_2, \dots, X_n$ be mutually stochasicated independent random variables having, respectively, the normal distributions $n(\mu_1,\sigma_1^2), n(\mu_2\sigma_2^2), \dots$ and $n(\mu_n,\sigma_n^2)$. The random variable $Y= k_1X_n+k_2x_2+ \dots + k_n, k_2 \dots, K_n$ are real constants, is normally distributed with mean $k_1\mu_1 + \dots + k_n\mu_n$ and variance $k^2{}_1\sigma^2{}_1+\dots+k^2n\sigma^2n$.

Proof :

Since Because $X_1,X_2, \dots, Xn$ re mutually stochastically independent the moment generating function of Y is given by

$M(t) = E\{exp[t(k_1X_1+ k_2X_2 +\dots+knXn)]\}$

$= E(e^{tk1X1})E(e^{tk2X2}) \dots E(e^{tKnXn})$

Now

$E(e^{tXi}) = exp(\mu_it+a_i^2t^2/2)$

for all real $t, i = 1,2 \dots, n$ Hence we have

$E(e^{tkixi}) = exp[\mu_i(k_it)+ \sigma i^2kit^2 /2]$,

That is, the moment generating function of Y is

$M(t) = {}^n\Pi_{i=1} exp[(k_i\mu_i)t + k^2{}_i\sigma_i^2)t^2/2]$

$= exp [(\{_1\Sigma^n k_{i\mu i})t+\Sigma^n{}_1 ki^2\sigma i^2)t^2/2]$

But this is the moment generating function of a distribution that is

$n({}^n\Sigma_1 k_i\mu_i, {}^n\Sigma_1 k_i^2 \sigma_i^2)$. Hence the proof.

Theorem 2:

Let $X_1, X_2, \dots X_n$ be mutually stochasicated independent variables that have respectively the chi-square distribution $X^2(r_1), X^2(r_2) \dots$, and $X^2(rn)$. then the random variable $Y=X_1+X_2+\dots+X_n$, has a chi-square distribution with $r_1+ \dots + rn$ degrees of freedom that is, Y is $X^2(r_1+..+\dots rn)$

Proof.

The moment generating func n(action of Y is

$M(t) = E\{exp[t((X_1+X_2+\dots+X_n))\}$ is

$= E(e^{tx1})E(e^{tx2})\dots E(e^{txn})$

Because X_1, X_2, ..., X_n are mutually stochastically independent since $E(e^{tx1}) = (1-2t)^{-r1/2}$ $t<1/2, t<1/2$ ....n

We have, $M(t) = (1-2t)^{-(r2+r2 \ldots\ldots +rn\ 1/2}$. $t<1/2$

But this is the moment generating function of a distribution that is $x^2(r_1+r_2+/\ldots+r_n)$. Accordingly Y has this ch-square distribution .

Also, let $X_1, X_2, \ldots, X_n$ be a random sample of size n from a distribution that is $n(\mu,\sigma^2)$ Thus, each of the randomk variable $(Xi-\mu^2 /\sigma 2$, $i=1,2,\ldots.n$ is $x^2(1)$. More over these n random variables are mutually stochastically independent . By date, the random variable $Y={}_1\Sigma^n\ [x(Xi-\mu)^2/\sigma^2$, $i=1,2..n$ is $X^2(n)$.

5.7 THE DISTRIBUTION OF $\overline{X}$ AND NS^2/σ^2

Let X_1 , $X_2, \ldots, X_n$ denote a random sample of size $n \geq 2$ from a distribution that is $n(\mu,\sigma^2)$. Here we discus about mean and the variance of this random sample that is the distribution of the two statistics

$X = {}^{\infty}\Sigma_1\ Xi/n$ and $S^2 = {}^n\Sigma_1\ (X_i-X)^2 / n$

The problem of the distribution of X. the mean of the sample is solved by the use of Theorem 1 if section 5.5. We have here , in the notation of the statement of that theorem $,\mu_1=m_2=\ldots.\mu_n=\ldots.\mu,\ \sigma_1^2 = \sigma_2^2 = \sigma^2_n=\sigma^2$ and $k_1=k_2=\ldots.=k_n=1/n$. Accordingly Y = X has a normal distribution with mean and variance given by

$${}^n\Sigma_1(1/n\mu) = \mu,\ \ {}_1\Sigma^n(1/n)^2\ \sigma^2\ ] = \sigma^2/n$$

respectively that is X is $n(\mu,\sigma^2/n)$

Example : Let X be the mean of a random sample of size 25 from a distribution that is n(75,100). Thus X is n(75,4) Then for instance,

$$Pr(71<X<79) = \frac{N(79-75)}{2} - \frac{N(71-75)}{2}$$

$= N(2)-N(-2) = 0.954$

We now take up the problem of the distribution of S^2 the variance of the random sample $X_1,...X_2...X_2$ from a distribution that is $n(\mu,\sigma^2$. Consider the joint distribution $Y_1=X_1$, $Y_2=X_2$, $Y_n=X_n$.

The corresponding transformation

$x_1 = ny_1-y_2.......y_n$

$x_2 = y_2$

... ...

$x_n = y_n$

has Jacobian n Since

$$\sum_1^n(x_1-\mu)^2 = \sum_1^n (x_i-\bar{x}+\bar{x}-\mu)^2 = \sum_1^n(x_1-\bar{x})^2 + n(\bar{x}-)^2$$

because $2(\bar{x}-\mu)\sum_1^n (x_i-\bar{x}) = 0$ the join p.d.f of $X_1,X_2,X_3....X_n$

can be written

$$(1\sqrt{2\pi\sigma})^n \exp\left[\frac{\sum(x_1-\bar{x})^2}{2\sigma^2} - \frac{n(\bar{x}-\mu)}{2\sigma^2}\right]$$

where x represents $(x1+x2+.......xn)/n$ and $-\infty<x1<\infty$, $i = 1, 2 , ... n$. Accordingly, with $y_1 =$

x, we find that the join p.d.f $Y_1,Y_2 ....Y_n$ is

$$n(1/\sqrt{2\pi\sigma})^n \exp\left[\frac{-(ny_1-y_2-......y_n-y_1)^2}{2\sigma^2}\right.$$

$$\frac{-2\sum^n (y_1-y_1)^2}{2\sigma^2} - \frac{n(y_1-\mu)^2}{2\sigma^2}$$

$-\infty<y_1<\infty$ $i=1,2,3......n$. The quotient of this join p.d.f and the p.d.f

$$\sqrt{n}/(\sqrt{2\pi\sigma})^{n-1}\exp[-n(y1-\mu)^2]$$

of $Y_1=X$ is the conditional p.d.f of $Y_2, Y_3 .... Y_n$ given $Y_1=y_1$

where $q=(ny_1-y_2-......y_n-y_1)^2+\sum(y_1-y_1)^2$. Since this is a join conditional p.d.f it must be, for all σ >0, that

$$\int_{-\infty}^{\infty}.........\int_{-\infty}^{\infty} \sqrt{n}\,(1/\sqrt{2\pi\sigma})^{n-1} \exp(-q/2\sigma^2)dy_2.....dy_n=1$$

Now consider

$nS_2 = \sum(Xt-X)^2$

$= (nY_1-Y_2-.......Yn)-yi)^2 + \Sigma(Yi-Y_1)^2 = Q$

The conditional moment generaing function of $nS^2/\sigma^2 = Q/\sigma^2$, given $Y_1=y_1$ is

$$E(e^{tQ}/\sigma^2/y_1) = {}_{-\infty}\int^{\infty} .... {}_{-\infty}\int^{\infty} \sqrt{n}\,(1/\sqrt{2\pi\sigma})^{n-1} \exp\left[-(1-2t)q\right] dy_2...dy_n \over 2\sigma^2$$

$$= \left(\frac{1}{1-2t}\right)^{n-1/2} {}_{-\infty}\int^{\infty} .... {}_{-\infty}\int^{\infty} \sqrt{n}\ \frac{(1-2t)^{(n-1)/2}}{2\pi\sigma^2}\ x\ \exp\left[-\frac{(1-2t)q}{2\sigma^2} dy_2.....dy_n\right]$$

Where $0<1-2t$, or $t<½$. However, this integral is exactly the same that of the conditional p.d.f of $Y_2,Y_3......Y_n$. given $Y_1=y_1$ with σ^2 replaced by $\sigma^2/(1-2t)>0$ and thus must equal 1. Hence the conditional moment generating function of nS^2/σ^2, given $Y_1 = y_1$ or equivalency $X = \bar{x}$, is

$$E(e^{tns2}/\sigma^2/\bar{x}) = (1-2t)-(n-1)/2, t<1/2$$

That the conditional distribution of nS^2/σ^2, given X=x, is $X^2(n-1)$. Moreover, since it is clear that this conditional distribution does not depend, upon $\bar{x}$, X and S^2 are stochachatically independent.

To summarize we have established, in this section, three important properties of X and S^2 when the sample arises from a distribution which is $n(\mu,\sigma^2)$:

(a) X is $n(\mu,\sigma^2/n)$

(b) nS^2/σ^2 is $X^2(n-1)$

(c) X and S^2 are stochastically independent.

Expectation of Functions of Random Variables

5.8 EXPECATIONS OF FUNCTIONS OF RANDOM VARIABLES

Theorem

Let $X_1,X_2.....X_n$ denote random variables that have means $\mu_1......\mu_n$ and variances σ^2_1 $.......\sigma^2n$. Let $p_{i,j}$ $i\neq j$ denote the correlation coefficient of Xi and Xj and let $k_1.....K_n$ denote real constants. The mean and the variance of the linear function

$$Y = {}^n\Sigma_i\ k_iX_i$$

are respectively

$$\mu_r = {}^n\ \Sigma_i\ k_i\mu_i$$

and $\sigma^2 y = {}^n\Sigma_1 k_i^2 \sigma_i^2 + 2\Sigma_{i<j} \Sigma k_i k_j \sigma_i \sigma_j$

Corollary Let $X_1, X_2 \ldots\ldots X_n$ denote the items of a random sample of the variance of $Y = {}^n\Sigma_1 k_i X_i$ are respectively $\sigma_r = ({}^n\Sigma_1 K_i^2)\ \mu$ and σ^2

Example 3

Let $X = {}^n\Sigma_1 X/n$ denote the mean of a random sample size from a distribution that has mean μ and variance σ^2. In accordance with the corollary, we have $\mu x = \mu\ {}^n\Sigma_1(1/n) = \mu$ and $\sigma^2 x = \sigma^2\ {}^n\Sigma_1(1/n)^2 = \sigma^2/n$. We have seen, in section 4.8 that if our sample is form a distribution that is n (μ, σ^2), then X is $n(\mu, \sigma^2/n)$. it is interesting that $\mu x = \mu$ and $\sigma = \sigma$ whether the sample is or not from a normal distribution.

5.9 LIMITING DISTRIBUTIONS

If X is the mean of a random sample $X_1, X_2, \ldots\ldots X_n$ from a distribution that has the p.d.f

$$f(x) = 1,\ 0<x<1,$$
$$= 0 \text{ elsewhere}$$

the moment generating function of X is given by [M(t/n)]n, where

$$M(t) = {}_0\int^1 e^{tx}\, dx = \frac{e^t - 1}{t},\ t \neq 0$$

$$= 1,\ t = 0$$

$$E(e^{t\bar{x}}) = \frac{(e^{t/n} - 1)}{(t/n)},\quad t \neq 0,$$

$$= 1,\ t = 0$$

$$Fn(\bar{x}) = {}_{-\infty}\int^{\bar{x}} \frac{1}{\sqrt{1/n}\sqrt{2\pi}} e^{-nw^2/2}\, dw$$

for the distribution function of the mean X_n of the random sample of the size n from a normal distribution with mean zero and variance 1

Definition

Let the distribution function F_n (y) of the random variable Y_n depend upon n, a positive integer. If F(y) is a distribution function and if lim $F_n(y) = F(y)$ for every point y at which F(y) is

$n \to \infty$

continuous, then the random variable Y_n is said to have a limiting distribution with distribution function F(y)

Example

Let Y_n denote the nth order statistic of a random sample $X_1, X_2 \ldots X_n$ from a distribution having p.d.f

$f(x) = 1/\theta,\ 0<\theta<\infty,$

$= 0$ elsewhere

The p.d.f of Y_n is

$g_n(y) = ny^{n-1}/\theta^n, \quad 0<y<\theta,$

$= 0$ elsewhere,

and the distribution function of Y_n is

$F_n(y) = 0, \ y<0$

$= \int_0^y \frac{nz^{n-1}}{\theta^n} dz = (y/\theta)^n, \ 0 \le y < \theta,$

$=1, \quad \theta \le y < \infty$

Then

$\lim_{n\to\infty} F_n(y) = 0, \ -\infty<y<\theta,$

$=1, \ \theta<y<\infty$

Now $F(y) = 0 \ -\infty<y<\theta$

$=1 \quad \theta \le y < \infty$

is a distribution function

Example

Let X_n have the distribution function

$$F_n(\bar{x}) = \int_{-\infty}^{\bar{x}} \frac{1}{\sqrt{1/n}\sqrt{2\pi}} e^{-nw^2/2} dw$$

If the change of variable $v = \sqrt{n}w$ is made we have

$Fn(\bar{x}) = \int_{-\infty}^{\sqrt{n}x} 1/\sqrt{2\pi}\,(e^{-v^2/2})\,dv$

Clearly,

$$\lim_{n \to \infty} Fn(x) = 0\ \bar{x}<0$$
$$= 1/2\ \bar{x}=0$$
$$= 1,\ \bar{x}>0$$

$F(x) = 0,\ \bar{x}<0$

$= 1,\ \bar{x}>0$

is a distribution function and $\lim_{x \to \infty} Fn(\bar{x}) = F(\bar{x})$ at every point of continuity of F(x).

Accordingly the random variable X_n has a limiting distribution with distribution function $F(\bar{x})$. Again this limiting distribution is degenerate and has all the probability at the one point $\bar{x} = 0$

Example

The fact that limiting distributions, if they exist cannot general be determined by taking the limit of p.d.f will now be illustrated let Xn have the p.d.f

$f(x) = 1,\ x = 2+1/n$

$= 0$ elsewhere

Clearly, $\lim_{n \to \infty} f_{n(x)=0}$ for all values of x. This may suggest that x_{n} is

$Fn(x) = 0 \quad x<2+1/n,]$

$= 1, \quad x \geq 2+1/n,$

and

$\lim_{n \to \infty} fn(x) = 0,\ x \leq 2$

$= 1,\ x \geq 2$

Since

$F(x) = 0, \quad x<2,$

$=1, \quad x\geq 2,$

is a distribution function, and since lim $fn(x) = F(x)$ at all points of continuity of F(x), there is a limiing distribution of X_n with distribution function F(x)

5.10 STOCHASTIC CONVERGENCE

Theorem

Let $F_n(y)$ denote the distribution function of a random variable Y_n whose distribution depends upon the positive integer n . Let c denote a constant which does not depend on n . The random variable Y_n converges stochastically to the constant c if and only if, for ever $\epsilon>0$, , the lim $Pr(|Yn-c|<\epsilon) =1$.

n->α

~~Proof. First assume that the ∈ >0, the~~

~~lim Pr(|Yn c)|<~~

—

Proof : Let

lim $Pr(|Yn-c|<\epsilon)=1$. for every

n->α

We have to prove that the random variable Y_n converges stochastically to the constant c. That is we have to prove that

lim $F_n(y) = 0, \quad y<c,$

n->α $\quad =1, \quad y>c.$

If the limit of $F_n(y)$ is indicated, then Y_n has a limiting distribution with distribution function

$F(y) = 0, y<c,$

$= 1, y\geq c.$

Now

$Pr(| yn-c|<\epsilon) = F_n [(c+\epsilon-]-F_n(c-\epsilon),$

where $F_n [(c+\epsilon\}-]$ is the left-hand limit of Fn(y) at $y=c+\epsilon$. Thus we have

$1 = \lim_{n\to\alpha} Pr(Y_n - c < \in) = \lim_{n\to\alpha} \leq F_n[(c+\in)-] - \lim_{n\to\alpha} Fn(c-\in)$

Because $0 \leq Fn(y) \leq 1$ for all values of y and for every positive integer n, it must be that

$\lim_{n\to\alpha} F_n(c-\in) = 0$, $\lim_{n\to\alpha} Fn[(c+\in)-] = 1$

Since this is true for every $\in > 0$, we have

$\lim_{n\to\alpha} F_n(y) = 0, y<c,$

$= 1, y>c,$

Now, we assume that

$\lim_{n\to\alpha} F_n(y) = 0 \; y<c,$

$1 = y>c.$

We are to prove that $\lim_{n\to\infty} Pr(|Yn - c| < \in) = 1$ for every $\in > 0$.

Because

$Pr(|Yn-c|<\in) = F_n[(c+\in)-] - F_n(c-\in)$,

and because it is given that $\lim_{n\to\alpha} F_n[(c+\in)-] = 1$,

$\lim_{n\to\alpha} F_n(c-\in) = 0)$,

for every $\in > 0$, we have the desired result. This completes the proof of the theorem.

That is this last limit is also a necessary and sufficient condtion for the stochastic Convergence of the random variable yn to the constant c

Example

Let X_n denote the mean of a random sample of size n from a distribution that has a mean μ and positive variance σ^2. Then the mean and variance of X_n are μ and σ^2/n. Consider for every fixed $\in > 0$, the probability

$\Pr(|X_n-\mu|\geq\epsilon) = \Pr(|X_n-\mu|\geq k\sigma/\sqrt{n})$, where $k = \epsilon\sqrt{n}/\sigma$. In accordance with the inequality of Chebyshev, this Probability is $\leq 1/k^2 = \sigma^2/n\epsilon^2$. So for every fixed $\epsilon>0$, we have

$$\lim_{n\to\infty} \Pr(|X_n-\mu|\geq\epsilon) \leq \lim \sigma^2/n\epsilon^2 = 0$$

Hence X_n converges stochastically to μ if σ^2 is finite

5.11 LIMITING MOMENT – GENERATING FUNCTIONS

Result:

Let the random variable Y_n have the distribution function $F_n(y)$ and th4e moment generating function $M(t;n)$ that exists for $-h<t<h$ for all n. If there exists a distribution functions $F(y)$, with corresponding moment generating function $M(t)$, defined for $|t| \leq h1,<h$, such the

$\lim_{n\to\infty} M(t;n) = M(t)$, then Y_n

has a limiting distribution with distribution function $F(y)$

Example 1

Let Y_n have a distribution that is $b(n,p)$. Suppose that the mean $\mu = np$ is the same for every n; that is $p=\mu/n$ where μ is a constant. We shall find the limiting distribution of the binomial distribution, when $p=\mu/n$, by finding the limit of $M(t;n)$. Now

$$M(t;n) = E(e^{tY_n}) = [(1-p)+pe^t]^n = \left[1+\frac{\mu(e^t-1)}{n}\right]^n$$

for all real values of t. Hence we have

$$\lim_{n\to\infty} M(t;n) = e^{\mu(e^t-1)}$$

for all real values of t. Since there exists a distribution, namely the poisson distribution with mean μ, that has this moment generating function $e^{\mu(e^t-1)}$

then in accordance with the theorem and under the conditions stated, it is seen that Y_n has a limiting poisson distribution with mean μ

Example 2

Let Z_n be $\chi^2(n)$. Then the moment generating function of Z_n is $(1-2t))^{-n/2}$, $t<1/2$. The mean and the variance of Z_n are respectively n and 2n. The limiting distribution of the random variable $Y_n=(Z_n-n)/\sqrt{2n}$ will be investigated. Now the moment generating function of Y_n is

$$M(t;n) = E\{\exp[t(Z_n-n)]\}$$

$$\sqrt{2n}$$

$$= e^{-tn/\sqrt{2n}} E(e^{tzn/\sqrt{2n}})$$

$$= \exp[-(t\sqrt{2/n})(n/2)](1-2t/\sqrt{2n})^{-n/2},\ t<\sqrt{2n}/2$$

This may be written in the form $M(t;n)=(e^{t\sqrt{2/n}}-t\sqrt{2/n}))^{-n/2}$, $t<\sqrt{4/2}$.

In accordance with Taylor's formula, there exists a number $\varepsilon(n)$, between 0 and $t\sqrt{2/n}$, such that

$$e^{t\sqrt{2/n}} = 1+t\sqrt{2/n}+1/2(t\sqrt{2/n})^2+_e\varepsilon(n)/6\ (t\sqrt{2/n})3$$

If this sum is substituted for $e^{t\sqrt{2/}}n$ in the last expression for $M(t;n)$, it is seen that

$$M(t;n) = (1-t2/n+\psi(n)/n)-x/2$$

where

$$\psi(n) = \frac{\sqrt{2}t^3 e^{\varepsilon n}}{3\sqrt{n}} - \frac{\sqrt{2}t^3}{\sqrt{n}} - \frac{2t^4 e^{\varepsilon(n)}}{3n}$$

Since $\varepsilon(n) \to 0$ as $n \to \infty$, then $\lim \psi(n)=0$ for every fixed value of t.

Also $\lim_{n\to\infty} M(t;n) = e^{t2/2}$

for all real values of t. That is the random variable $Y_n = (Z_n-n)/\sqrt{2n}$ has a limiting normal distribution with mean zero and variance 1.

5. 12 THE CENTRAL LIMIT THEOREM

Statement : Let $X_1, X_2......X_n$ denote the items of random sample from a distribution that has mean μ and positive variance σ^2. Then the random variable $Y_n = (\sum_{1}^{n} X_i t - n\mu)/\sqrt{n}\sigma = \sqrt{n}(X_n-\mu)/\sigma$ has a limiting distribution that is normal with mean zero and variance 1

Proof : We assume the existence of the moment generating function $M(t) = E(e^{tx})$, $-h<t<h$, of the distribution.

The function

$$m(t)=E[e^{t(x-\mu)}] = e^{-\mu t M(t)}$$

also exists for $-h<t<h$. Since, m(t) is the moment generating function for $X-\mu$, it must follow that $m(0)=1$, $m^1(0) = E(X-\mu)=0$ and $m^{ll}(0)=E[(X-\mu)^2] = \sigma^2$

By Taylor's formula, there exists a number ξ between 0 and t such that

$$m(t)=m(0) + M1(0)t+m^{ll}\varepsilon)t^2/2$$

$$= 1+m''(\varepsilon)t^2/2$$

If $\sigma^2t^2/2$ is added and subtracted, then

$$m(t) = 1+\sigma^2t^2/2+[m''(\varepsilon)-\sigma^2]t^2/2$$

Now consider M(t;n), where

$$M(t;n) = E[\exp(\frac{t\sum X_i - n\mu)}{\sigma\sqrt{n}}]$$

$$= E[\exp(\frac{tX_1-\mu)}{\sigma\sqrt{n}} \exp(\frac{tX_2-\mu)}{\sigma\sqrt{n}} ....\exp \frac{(tX_n-\mu)}{\sigma\sqrt{n}}]$$

$$= E[\exp \frac{(tX1-\mu)}{\sigma\sqrt{n}}]......E[\exp(\frac{tXn-\mu)}{\sigma\sqrt{n}}]$$

$$= \{E[\exp(\frac{tX-\mu)}{\sigma\sqrt{n}}]\}n$$

$$= [m(t/\sigma\sqrt{n}\}]n, \qquad -h<t/\sigma\sqrt{n}<h$$

In m(t) replace t by $t/\sigma\sqrt{n}$ to obtain

$m(t/\sigma\sqrt{n})=1+t^2/2n+[m''(\varepsilon)-\sigma^2]t^2/2n\sigma^2$

where now ε is between o and $t/\sigma\sqrt{n}$ with $-h\sigma\sqrt{n}<t<h\sigma\sqrt{n}$

Accordingly

$$M(t;n)=\left\{1+\frac{t^2/2n+[m''(\varepsilon)-\sigma^2]t^2}{2n\,\sigma^2}\right\}n$$

Since $m''(t)$ is continuous at $t=0$ and since $\varepsilon\to0$ as $n\to\infty$, we have

$\lim [m''(\varepsilon)-\sigma^2]=0$

Thus, $\lim M(t;n) = et^2/2$

for all real values of t. This proves that the random variable $Y_n=\sqrt{n}(X_n-\mu)/\sigma$ has a limiting normal distribution with mean zero and variance 1.

Result

Let Fn(u) denote the distribution function of a random variable U_n whose distribution depends upon the positive integer n. let U_n converge stochastically to the constant $c\neq0$. The random variable U_n/c converges stochastically to 1.

Theorem

Let $F_n(u)$ denote the distribution function of a random variable U_n whose distribution depends upon the positive integer n. Further, let U_n converge stochastically to the positive constant c and let Pr $(Un<0)$ = 0 for every n. The random varicable $\sqrt{U_n}$ converges stochasticcally to $\sqrt{c}$.

Proof. We are given that the lim Pr $|\sqrt{U_n}-\sqrt{c}|\geq\varepsilon) = 0$ for every $\varepsilon>0$. We have to prove that the lim Pr $(|un - c| \geq \epsilon 1)=0$ for every $\epsilon^1>0$. Now the probability,

$\Pr(|U_n-c|\geq\varepsilon) = \Pr[|\sqrt{U_n}-\sqrt{c})(\sqrt{U_n}+\sqrt{c})|\geq\varepsilon]$

$=\Pr(|\sqrt{Un}-\sqrt{c}|\geq\varepsilon / \sqrt{Un}+\sqrt{c})$

$\geq\Pr(|\sqrt{Un}-\sqrt{c}|\geq\varepsilon\sqrt{c})\geq0.$

if we let $\varepsilon' = \varepsilon/\sqrt{c}$, and if we take the limit, as n becomes infinite, we have

$$0=\lim_{n\to\infty} \Pr(|Un-c|\geq\varepsilon)\geq \lim \Pr(|\sqrt{Un}-\sqrt{c}|\geq\varepsilon^1)=0$$

for every $\varepsilon^1>0$. This completes the proof.

Hence the proof.

EXERCISES

(1) Show that

$$S^2 = \frac{1}{n} \sum_1^n (X_i-X)^2 = \frac{1}{n} \sum_1^n X_i^2 - X^2,$$

Where $X = \sum_1^n X_i/n$.

(2). Find the probability that exactly four items of a random sample of size 5 from the distribution having p.d.f . $f(x)=(x+1)/2$, $-1<x<1$, zero else where exceed zero.

(3). Let X_1,X_2 be a random sample from the distribution having p.d.f. $f(x)=2x$, $0<x<1$, zero elsewhere. Find $Pr(X_1/X_2 \le 1/2)$.

(4) If the sample size is n=2, find the constant c so that $s^2=c(X_1-X_2)^2$.

(5). If $x_i = i$, $i=1,2,\ldots,n$, compute the values of $x=\sum xi/n$ and

$s_2=\sum(x1-x)2/n$.

(6) Let $yi=a+bxi$, $i=1,2,\ldots,n$, where a and b are constants. Find $y=\sum y_\alpha/n$ and $s^2y=\sum (yi-y)2/n$ in terms of a, b, $x=\sum xi/n$ and $s^2_x=\sum(x_i-x)^2/n$.

(7). Let X have a p.d.f. $f(x)=1/3$, $x=1,2,3$, zero elsewhere. Find the p.d.f. of $Y=2X+1$.

(8). If $f(x_1,x_2)=(2/3)x_1+x_2(1/3)^2 - x_1-x_2,(x_1-x_2)=(0,0),(0,1),(1,0),(1,1)$

zero elsewhere, is the joint p.d.f $x_1=x_2$ find the joint p.d.f. og $y_1=x_1-x_2$ and $y_2=x_1+x_2$

(9). Let X have the p.d.f $f(x) = (½)x$, $x=1,2,3\ldots$. zero elsewhere. Find the p.d.f of $Y=x_3$.

(10). Let X have the p.d.f $f(x) = x^2/9$, $0<x<3$, zero elsewhere. Find the p.d.f of $y=X^3$

(11). If the p.d.f of X is $f(x)=2x\,e^{-x2}$, $0<x<\infty$, zero elsewhere determine the p.d.f of $Y= X^2$.

(12). Let X^1, X^2 be a random sample from the normal distributes n(0,1). Show that the marginal p.d.f of $Y^1 = X_1/X_2$ is the Cauchy p.d.f. $g_1(y_1)=1/\pi(1+y_1^2)$, $-\infty<y_1<\infty$

(13). Let the stochastically independent random variables X_1 and X_2 have the same p.d.f $f(x)=1/6$, $x = 1,2,3,4,5,6$ zero elsewhere . Find the p.d.f of $Y=X_1+X_2$. Note under appropriate assumptions 11that Y may be interpreted as the sum of the spots that appear when two dice are cast.

(14). Let X_1 and X_2 be stochastically independent with normal distribution n(6,1) and n(7,1), respectively. Find $Pr(X_1>X_2)$. Hint. Write $Pr(X_1>X_2) = Pr(X_1-X_2>0)$ and determine the distribution of X_1-X_2.

(15). Let $X_1,X_2,...X_n$ denote n mutually stochastically independent random variables with the moment generating functions $M_1(t),M_2(t),....M_n(t)$, respectively.

(a) Show that $Y=k_1X_1+k_2X_2+....+k_nX_n$, where $k_1,k_2,...k_n$ are real constants, has the moment generating function M(t)Π Mi(kit).

(b) If each ki =1 and if Xi is poission with mean μ,i =1,2,....n prove that Y is poisson with mean μ1+...+μn.

(16). Let X_n denote the mean of a random sample of size n from distribution that is $n(\mu,o^2)$ Find the limiting the distribution of Xn.

(17). Let X_n have a gamma distribution with parameter α=n and β and β is not a function of n. Let Y_n = Xn/n. Find the limiting distribution of Yn.

(18). Let Z_n be χ2(n) and let Wn=Zn/n2./ Find the limiting distribution of Wn.

(19). Let X be χ2(50). Approximate Pr(40<X<60).

QUESTION PAPER PATTERN

PART - A (5 x 5 = 25 marks)

Answer FIVE Question out of EIGHT Question

PART - B (5 x 15 = 75 marks)

Answer FIVE Question out of EIGHT Question