SEMESTER : II
CORE COURSE : VIII

Inst Hour	: 6
Credit	: 4
Code	: 18KP2M08

PROBABILITY THEORY

UNIT – I

Fields and σ Fields – Class of events – Functions and Inverse functions – Random variables – Limits of random variables.

Chapter 1 and 2 Omit (1.1 & 1.2)

UNIT – II

Probability Space: Definition of Probability – Some simple properties – Discrete Probability space – General probability space – Induced probability space.

Chapter 3 (Omit 3.6)

UNIT – III

Distribution functions: Distribution functions of a random variable – Decomposition of distributive functions – Distributive functions of vector random variables – Correspondence theorem.

Chapter 4

UNIT – IV

Expectation and Moments: Definition of Expectation – Properties of expectation – Moments. Inequalities.

Chapter 5

UNIT – V

Convergence of Random Variables: Convergence in Probability – Convergence almost surely – Convergence in distribution – Convergence in the r^{th} mean – Convergence theorems for Expectations.

Chapter 6 (6.1 to 6.5)

TEXT BOOK

B.R.Bhat (2007), MODERN PROBABILITY THEORY, 3rd Edition, New Age International private limited, New Delhi.

REFERENCES

1. Chandra T.K. and Chatterjee D.(2003), A first Course in Probability, 2nd Edition, Narosa Publishing House, New Delhi.
2. Kailai Chung and Farid Aitsahlia, Elementary Probability, Springer Verlag 2003, New York.
3. Marek Capinski and Thomasz zastawniak(2003), Probability through problems, Springer Verlag, NewYork.
4. Sharma.T.K(2005), A Text Book of probability and theoretical distribution, Discovery publishing house, New Delhi.

Question Pattern

Section A : 10 x 2 = 20 Marks, 2 Questions from each Unit.
Section B : 5 x 5 = 25 Marks, EITHER OR (a or b) Pattern, One question from each Unit.
Section C : 3 x 10 = 30 Marks, 3 out of 5, One Question from each Unit.

I M.Sc. MATHEMATICS

TITLE OF THE PAPER : PROBABILITY THEORY

PAPER CODE : 18KP2M08

UNIT - I

Definition - Closure

Suppose $\mathcal{A}$ is a class of sets. $\mathcal{A}$ is said to be closed under the following operations:

i) If $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$ then $\mathcal{A}$ is closed under complementation.

ii) If $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$ then $\mathcal{A}$ is closed under unions.

iii) If $A_1, A_2, \dots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$, $\forall n < \infty$ then $\mathcal{A}$ is closed under finite unions.

iv) If $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$ then $\mathcal{A}$ is closed under intersection.

v) If $A_1, A_2, \dots, A_n \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{n} A_i \in \mathcal{A}$, $\forall n < \infty$ then $\mathcal{A}$ is closed under finite intersections.

Example-1:

Let $\mathcal{C}$ be the class of all intervals of the form $(x, \infty) \in R$. Considered as subset of the real line. Show that $\mathcal{C}$ is not closed under complementation.

Proof:

$\mathcal{C}$ is closed under unions and intersections.

Since $(x, \infty) \cup (y, \infty) = (u, \infty)$ and $(x, \infty) \cap (y, \infty) = (v, \infty)$

where $u = \min(x, y)$ and $v = \max(x, y)$. Hence $\mathcal{C}$ is closed under finite unions and intersections. But $(x, \infty)^c = (-\infty, x) \notin \mathcal{C}$.

Hence $\mathcal{C}$ is not closed under complementation.

Example-2:

For the class $\mathcal{C}$ of all intervals of the form (a, b), $a, b \in R$ $(a < b)$ and ϕ. Show that $\mathcal{C}$ is not closed under complementation.

Proof:

(a, b), $a, b \in R$ $(a < b)$ and ϕ

$$\begin{aligned}
\rightarrow (a,b) \cap (c,d) &= \phi, \text{ if } a<b<c<d \text{ or } c<d<a<b \\
&= (c,b) \text{ if } a<c<b<d \\
&= (a,d) \text{ if } c<a<d<b \\
&= (c,d) \text{ if } a<c<d<b \\
&= (a,b) \text{ if } c<a<b<d
\end{aligned}$$

Thus, $\mathcal{C}$ is closed under finite intersections. But it is not closed under complementations or unions because $(a,b)^c = (-\infty, a] \cup [b, \infty) \notin \mathcal{C}$ and $(a,b) \cup (c,d)$ is not an interval if $a < b < c < d$ or $c < d < a < b$.

Definition - Field:

A nonempty class of sets $\mathcal{A}$, closed under complementation and finite intersections is called a field (or algebra).

Lemma 1:

A field is closed under finite unions. Conversely, a class closed under complementations and finite unions is a field.

Proof:

Suppose $\mathcal{A}$ is a field. Then

i) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

ii) $A_1, A_2, \ldots, A_n \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{n} A_i \in \mathcal{A}$. But $A_1, A_2, \ldots, A_n \in \mathcal{A}$
$\Rightarrow A_1^c, A_2^c, \ldots, A_n^c \in \mathcal{A} \Rightarrow \bigcap_{i=1}^{n} A_i^c \in \mathcal{A} \Rightarrow \left(\bigcap_{i=1}^{n} A_i^c\right)^c \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

Hence $\mathcal{A}$ is closed under finite unions also.

Conversely, suppose $\mathcal{A}$ is a class such that $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.
(iii) $A_1, A_2, \ldots, A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$. Then proceeding as above and using De-Morgan Rule, it can be proved that $\mathcal{A}$ is closed under finite intersections also.

Lemma 2:

Every field contains the empty set ϕ and the whole space Ω.

Proof:

The class containing only ϕ and Ω is a field. It is the smallest field and is contained in every other field. It is called the degenerate or trivial field. The power set consisting of every subset of a finite Ω is also a field and is the largest field. If a field contains A it has to contain A^c, and hence contains the class $\{A, A^c, \phi, \Omega\}$.

But this class is a field, contained in every field containing A. Therefore it is the smallest field containing A.

DEFINITION - MINIMAL FIELD:

Let $\mathcal{E}$ be an arbitrary class of sets. The smallest field containing $\mathcal{E}$ is called the minimal field containing $\mathcal{E}$ or the field generated by $\mathcal{E}$.

THEOREM:

The intersection of arbitrary number of fields is a field.

Proof:

Let $\mathcal{E}_i$ $(i \in I)$ be the fields indexed by points of a nonempty set I which may be finite, countable or uncountable. Let $\mathcal{E}_0 = \bigcap_{i \in I} \mathcal{E}_i$ the class of all sets common to all $\mathcal{E}_i$. Since each one of the $\mathcal{E}_i$ contains ϕ and Ω, $\mathcal{E}_0$ contains ϕ and Ω. Suppose $A \in \mathcal{E}_0$, then $A \in \mathcal{E}_i$, $\forall i \in I$. Hence $A^c \in \mathcal{E}_i$ for all i. Thus $A^c \in \mathcal{E}_0$ and $\mathcal{E}_0$ is closed under complementation.

If $A_k \in \mathcal{E}_0$, $(k = 1, 2, \ldots, n)$ then $A_k \in \mathcal{E}_i$, $\forall i \in I$. Then $\bigcap_{k=1}^{n} A_k \in \mathcal{E}_i$ for all $i \in I$ and hence it belongs to $\bigcap_{i=1}^{n} \mathcal{E}_i = \mathcal{E}_0 \Rightarrow \mathcal{E}_0$ is closed under finite intersections. Hence $\mathcal{E}_0$ is a field.

DEFINITION - PARTITION:

Let $\{A_i\}$ be mutually exclusive and exhaustive. Then $\{A_i\}$ is called a partition of Ω.

DEFINITION:- σ-FIELD OR σ-ALGEBRA:

A nonempty class of sets which is closed under complementations and countable unions (or countable intersections) is called a σ-field or σ-algebra.

THEOREM:

The intersection of an arbitrary number of σ-fields is a σ-field.

Proof:

For a given class $\mathcal{E}$ the minimal σ-field containing the class $\mathcal{E}$ will be denoted by $\sigma(\mathcal{E})$. It is the intersection of all σ-field containing $\mathcal{E}$. It is also called the σ-field generated by $\mathcal{E}$. If $\mathcal{E}$ is finite, the minimal field $F(\mathcal{E})$ containing $\mathcal{E}$, coincides with the minimal σ-field $\sigma(\mathcal{E})$ containing $\mathcal{E}$.

To obtain $\sigma(\mathcal{E})$ from a given class $\mathcal{E}$:

(i) obtain $\mathcal{E}_1 = \{\phi, \Omega, A, A^c \ni A \in \mathcal{E}$ or $A^c \in \mathcal{E}\}$ where $A \subset \Omega$. Evidently $\mathcal{E}_1$ is closed under complementation and contains $\mathcal{E}$.

(ii) Obtain the class $\mathcal{E}_2$ containing $\bigcap_{k=1}^{\infty} B_k$, where $B_k \in \mathcal{E}_1$, $k=1,2,\ldots,n$, B_k's and n being arbitrary. Now $\mathcal{E}_2$ is closed under finite intersections and not complementation.

(iii) obtain $\mathcal{E}_3$, the class of all finite unions of pair-wise disjoint sub-sets belonging to $\mathcal{E}_2$. Since they also contain complements, $\mathcal{E}_3$ is a field and is the minimal field containing $\mathcal{E}$.

Therefore $\mathcal{E}$ is a σ-field and the intersections is also a σ-field.

Lemma:

Let $\mathcal{E}_1$ be the class of all intervals of the form $(a,b)\,(a<b)$, $a, b \in R$, but arbitrary., Then $\sigma(\mathcal{E}_1) = \mathcal{B}$.

Proof:

$(a,b) \in \mathcal{B}$ for all a, b. Hence $\mathcal{E}_1 \subset \mathcal{B}$. By definition of minimal σ-field $\sigma(\mathcal{E}_1) \subset \mathcal{B}$.

To prove inclusion in the reverse direction, by the definition of $\sigma(\mathcal{E}_1)$
$\bigcup_{n=1}^{\infty}(-n,x) \in \sigma(\mathcal{E}_1)\, \forall x \Rightarrow (-\infty,x) \in \sigma(\mathcal{E}_1). \forall x \Rightarrow \mathcal{E} \subset \sigma(\mathcal{E}_1)$ where $\mathcal{E}$ is known.
Hence $\sigma(\mathcal{E}) \subset \sigma(\mathcal{E}_1)$. But $\sigma(\mathcal{E}) = \mathcal{B}$ and therefore $\sigma(\mathcal{E}) = \sigma(\mathcal{E}) = \mathcal{B}$.

DEFINITION – MONOTONE FIELD:

A field $\mathcal{F}$ is said to be Monotone Field, if it is closed under monotone operations., if $\lim A_n \in \mathcal{F}$ whenever $\{A_n\}$ is a monotone sequence of sets of $\mathcal{F}$. (ie) $A_n \in \mathcal{F}$, $A_n \uparrow A \Rightarrow A \in \mathcal{F}$ and $A_n \in \mathcal{F}$, $A_n \downarrow A \Rightarrow A \in \mathcal{F}$.

THEOREM:

A σ-field is a monotone field and conversely.

Proof:

Let $\mathcal{F}$ be a σ-field and $A_n \in \mathcal{F}$. If $A_n \uparrow A$, then $A = \bigcup_n A_n$ is a countable union of sets of F. Hence $A \in F$. Similarly if $A_n \downarrow A$, $A \in F$. Hence F is a monotone field.

Conversely, let F be a monotone field and let $A_1, A_2, \ldots, A_n$ be sets belonging to F. Then $\bigcup_{k=1}^{n} A_k$ and $\bigcap_{k=1}^{n} A_k$ belong to F, since F is a field. These are monotone sequences whose limits $\bigcup_{k=1}^{\infty} A_k$ and $\bigcap_{k=1}^{\infty} A_k$ must belong to F. Thus F is a σ-field.

DEFINITION - FUNCTION:

A function X on a space Ω to a space Ω' assigns to each point $\omega \in \Omega$ a unique point in Ω' denoted by $X(\omega)$. $X(\omega)$ is the image of the argument ω under X. It is also called the value of X at ω.

The space Ω is called the domain of X and Ω' is called the range. The set $\Omega'' = [X(\omega) : \omega \in \Omega]$, which is a subset of Ω', is called the strict range of X.

If the range space is the real line R or its subset, the function is said to be a numerical or real-valued function.

DEFINITION - INVERSE FUNCTION:

With every point function X, we associate a set function X^{-1} whose domain is a class $\mathcal{B}'$ of subsets of Ω' and whose range is a class $\mathcal{B}$ of subsets of Ω. The function X^{-1} is called inverse function (or mapping) of X.

$$X(B) = [X(\omega) : \omega \in B],\ B \subset \Omega.$$

$$X^{-1}(\mathcal{B}') = [X^{-1}(B') : B' \in \mathcal{B}'].$$

Lemma:

Inverse mapping preserves all set relations:

Proof:

(i) Let $B \subset C \subset \Omega'$. Then $X^{-1}(B) = [\omega : X(\omega) \in B] \subset [\omega : X(\omega) \in C] = X^{-1}(C)$

(ii) Let $B_k \subset \Omega'$. Then

$$\omega \in X^{-1}(\cap B_k) \Leftrightarrow X(\omega) \in \cap B_k \Leftrightarrow X(\omega) \in B_k,\ \forall k.$$
$$\Leftrightarrow \omega \in X^{-1}(B_k),\ \forall k \Leftrightarrow \omega \in \cap X^{-1}(B_k)$$

Hence $X^{-1}(\bigcap_k B_k) = \bigcap_k X^{-1}(B_k)$.

Similarly we can prove that.

$$X^{-1}(\cup B_k) = \bigcup_k X^{-1}(B_k).$$

(iii) $\omega \in X^{-1}(B^c) \Leftrightarrow X(\omega) \in B^c \Leftrightarrow X(\omega) \notin B \Leftrightarrow \omega \notin X^{-1}(B)$.

Thus $X^{-1}(B^c) = (X^{-1}(B))^c$.

DEFINITION - MEASURABLE FUNCTION:

If $X^{-1}(B) \in \mathcal{A}$ for all Borel sets $B \in \mathcal{B}$, X is said to be a $\mathcal{A}$ measurable function, or a function measurable with respect to $\mathcal{A}$.

If Ω is also the real line R or its subset and if X is measurable with respect to the Borel field $\mathcal{B}$ on the domain then X is called the Borel function.

DEFINITION - SIMPLE FUNCTION:

A finite linear combination of indicators of sets is called a simple function.

LEMMA:

Any simple function X can be written in the form $X = \sum_{k=1}^{n} x_k I A_k$, where x_k's are distinct numerical constants and $\{A_1, A_2, ---, A_n\}$ is a partition of Ω.

Proof:

Let X be a linear function of I_{B_1} and I_{B_2} only.

$$X = C_1 I_{B_1} + C_2 I_{B_2}$$

Then $X(\omega) = 0$ if $\omega \notin B_1$ and $\omega \notin B_2$. (ie) $\omega \in B_1^c B_2^c = A_1$,
$= C_1$, if $\omega \in B_1$ and $\omega \notin B_2$. (ie) $\omega \in B_1 B_2^c = A_2$,
$= C_2$, if $\omega \notin B_1$ and $\omega \in B_2$. (ie) $\omega \in B_1^c B_2 = A_3$.
$= C_1 + C_2$, if $\omega \in B_1$ and $\omega \in B_2$. (ie) $\omega \in B_1 B_2 = A_4$.

Hence $X(\omega) = x_i$ for $\omega \in A_i$, where $x_1 = 0$, $x_2 = C_1$, $x_3 = C_2$ and $x_4 = C_1 + C_2$. Moreover $\{A_1, A_2, A_3, A_4\}$ forms a partition of Ω.

Suppose in general X is a linear combination of m indicators, $X = \sum_{i=1}^{m} C_i I_{B_i}$. Then $X(\omega) = C_{j_1} + C_{j_2} + \cdots + C_{j_k}$ if $\omega \in B_{j_1} \cap B_{j_2} \cap \cdots \cap B_{j_k} \cap B_{j_{k+1}}^c \cdots \cap B_{j_m}^c = A_i$ (say) where $(j_1, j_2, \ldots, j_m)$ is a permutation of $(1, 2, \ldots, m)$ and $k = 0, 1, \ldots, m$. Thus X can have atmost 2^m values which are distinct. The sets A_i being disjoint, $\{A_i\}$ forms a partition of Ω.

Thus, a simple function takes only a finite number of values $x_1, x_2, \ldots, x_n$. The class $\{A_i\}$ is the partition on which X takes distinct values. Thus X is a simple function of the form $\sum_{k=1}^{n} x_k I_{A_k}$ with distinct x_k's and A_k's disjoint forming a partition of Ω.

LEMMA:

The σ-field induced by a simple function $X = \sum_{k=1}^{n} x_k I_{A_k}$ is the minimal σ-field containing the partition $(A_1, A_2, \ldots, A_n)$

Proof:

Let B be any Borel subset of the real line containing $[x_{j_1}, x_{j_2}, \ldots, x_{j_k}] \subset [x_1, x_2, \ldots, x_n]$.

Then $X(\omega) \in B$ iff $X(\omega) = x_{j_1}$ or x_{j_2} or $\ldots$ x_{j_k}.

But $X(\omega) = x_j$, if $\omega \in A_j$. Hence

$$[\omega : X(\omega) \in B] = X^{-1}(B) = \sum_{i=1}^{k} A_{j_i} \in \sigma\{A_1, A_2, \ldots, A_n\}.$$

Since the sets of $\sigma\{A_i\}$ are sums of sets of $\{A_i\}$.

DEFINITION - FUNCTION OF A FUNCTION:

If X is a function or a mapping from Ω to Ω' and X' is a function from Ω' to Ω'' then $X'(X(\omega))$ is a function from Ω to Ω'', denoted by $X'X$ or $X'(X)$. It is said to be a function of a function or a composition of two functions X and X'. Its inverse $(X'X)^{-1}$ is a function on the subset of Ω'' to the subsets of Ω such that for any $B \subset \Omega''$.

$$\begin{aligned}(X'X)^{-1}(B) &= [\omega : X'(X(\omega)) \in B] \\ &= [\omega : X(\omega) \in X'^{-1}(B)] \\ &= X^{-1}(X'^{-1}(B)).\end{aligned}$$

Thus, $(X'X)^{-1} = X^{-1}X'^{-1}$.

LEMMA:

Borel function of a $\mathcal{A}$-measurable function X is a $\mathcal{A}$-measurable and induces a sub-σ-field of that induced by X.

Proof:

If $f(X)$ is the Borel function of X and B is any Borel set. $[f(X)\in B] = [X \in f^{-1}(B)]$. But if B is a Borel set, $f^{-1}(B)$ is a Borel set B' and hence, $[X\in f^{-1}(B)] = [\omega : X(\omega)\in B'] \in X^{-1}(\mathcal{B}) \subset \mathcal{A}$. Thus $f(X)$ is $\mathcal{A}$ measurable and induces a sub-σ-field of $X^{-1}(\mathcal{B})$ which is the σ-field induced by X.

DEFINITION - RANDOM VARIABLE:

A real valued, $\mathcal{A}$-measurable function defined on Ω is called a random variable (r.v).

LEMMA:

X is a r.v iff $X^{-1}(\mathcal{E})\subset \mathcal{A}$, where $\mathcal{E}$ is any class of subsets of R, which generates $\mathcal{B}$.

Proof:

We have to prove that, $X^{-1}(\mathcal{E})\subset\mathcal{A} \Leftrightarrow X^{-1}(\mathcal{B})\subset\mathcal{A}$.
Since $\mathcal{E}\subset\mathcal{B}$ and $X^{-1}(\mathcal{B})\subset\mathcal{A}$, $X^{-1}(\mathcal{E})\subset\mathcal{A}$. To prove the converse since $\mathcal{A}$ is a σ-field and $X^{-1}(\mathcal{E})\subset\mathcal{A} \Leftrightarrow \sigma(X^{-1}(\mathcal{E}))\subset\mathcal{A}$
$\Leftrightarrow X^{-1}(\sigma(\mathcal{E}))\subset\mathcal{A} \Leftrightarrow X^{-1}(\mathcal{B})\subset\mathcal{A}$.

LEMMA:

$Z^{-1}(\mathcal{B}_2)$ is the smallest σ-field with respect to which X and Y are measurable, $Z^{-1}(\mathcal{B}_2) = \sigma[X^{-1}(\mathcal{B})\cup Y^{-1}(\mathcal{B})]$.

Proof:

Consider the rectangle,

$$B = [(x,y) : a<x<b,\ c<y<d],\ a,b,c,d\in R.$$

$$Z^{-1}(B) = [\omega : a< X(\omega)<b,\ c<Y(\omega)<d]$$
$$= [\omega : a<X(\omega)<b]\cap[\omega : c<Y(\omega)<d]$$
$$Z^{-1}(B) = X^{-1}(a,b)\cap Y^{-1}(c,d)$$

Taking $c=-\infty$, $d=+\infty$, we have $Y^{-1}(c,d) = Y^{-1}(R) = \Omega$.
Hence the class $[Z^{-1}(B) : B \text{ a rectangle}] \supset [X^{-1}(a,b) : a<b,\ a,b\in R]$
But $\sigma[X^{-1}(a,b), a,b\in R] = X^{-1}(\mathcal{B})$, the σ-field induced by X and

$\sigma\{Z^{-1}(B), B \text{ a rectangle}\} = Z^{-1}(\mathcal{B}_2)$, the σ-field induced by Z. $Z^{-1}(\mathcal{B}_2) \supset X^{-1}(\mathcal{B})$. Similarly by taking $a=-\infty$ and $b=+\infty$ and Continue as above, $Z^{-1}(\mathcal{B}_2) \supset Y^{-1}(\mathcal{B})$. Therefore $Z^{-1}(\mathcal{B}_2) \supset X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})$ and since $Z^{-1}(\mathcal{B}_2)$ is a σ-field. $Z^{-1}(\mathcal{B}_2) \supset \sigma[X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})]$. $\longrightarrow$ ①

To prove the reverse inclusion,

$$Z^{-1}(B) = X^{-1}(a,b) \cap Y^{-1}(c,d) \in \sigma[X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})] \longrightarrow ②$$

Since the R.H.S of ① is a σ-field and since

$\{Z^{-1}(B), B \text{ a rectangle}\} \subset \sigma\{X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})\}$

$\sigma[Z^{-1}(B), B \text{ a rectangle}] \subset \sigma[X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})]$

By $X^{-1}(\bigcap_k B_k) = \bigcap_k X^{-1}(B_k)$, $Z^{-1}(\mathcal{B}_2) \subset \sigma[X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})] \longrightarrow ③$

From ①&③ $Z^{-1}(\mathcal{B}_2) = \sigma[X^{-1}(\mathcal{B}) \cup Y^{-1}(\mathcal{B})]$.

DEFINITION – TAIL EVENTS:

The σ-field T contained in $\mathcal{E}_n$, for all n, is called the tail σ-field of $\{X_n\}$. The events of T are called tail events. A function (r.v) which induces a σ-field which is a sub-σ-field of T is called a tail-measurable function (r.v).

LIMITS OF RANDOM VARIABLES:

Let $\{X_n\}$ be a sequence of r.v.'s on $(\Omega, \mathcal{A})$. Let us define functions Y_n and Z_n on Ω as follows:

$Y_n(\omega) = \inf_{k \geq n} X_n(\omega)$ = greatest lower bound of $X_k(\omega)$ for $k \geq n$

$Z_n(\omega) = \sup_{k \geq n} X_k(\omega)$ = least upper bound of $X_k(\omega)$ for $k \geq n$.

As $n \to \infty$, $Y_n(\omega) \uparrow \sup Y_n(\omega) = \lim\inf X_n(\omega)$ $\underline{\lim} X_n(\omega)$

$Z_n(\omega) \downarrow \inf Z_n(\omega) = \lim\sup X_n(\omega)$ $\overline{\lim} X_n(\omega)$

These functions $\underline{\lim} X_n(\omega)$ and $\overline{\lim} X_n(\omega)$ will exist for all $\omega \in \Omega$. Moreover $\overline{\lim} X_n(\omega) \geq \underline{\lim} X_n(\omega)$, $\forall \omega$.

DEFINITION – SET OF CONVERGENCE:

$\{X_n\}$ Converges to X on A. $X_n(\omega) \to X(\omega)$, $\omega \in A$, if $\overline{\lim} X_n(\omega) = \underline{\lim} X_n(\omega) = X(\omega) < \infty$, $\forall \omega \in A$. The set A for all ω for which $X_n(\omega) \to X(\omega)$ is called the set of Convergence of the sequence $\{X_n\}$.

LEMMA:

If $X \geq 0$, there exists a monotone increasing sequence of non-negative simple functions $\{X_n\}$ converging to X.

Proof: Define

$$X_n(\omega) = k2^{-n}, \text{ if } k2^{-n} \leq X(\omega) \leq (k+1)2^{-n}, \ k = 0, 1, \ldots, n2^n - 1).$$

$$= n, \text{ if } X(\omega) \geq n. \text{ So that}$$

$$X_n = \sum_{k=0}^{n2^n-1} k2^{-n} I[k2^{-n} \leq X < (k+1)2^{-n}] + nI[X \geq n] \text{ is a simple r.v.}$$

Then $X_n(\omega) \geq 0$ for all n and ω. Moreover

$$X_{n+1}(\omega) = 2k2^{-(n+1)}, \text{ if } 2k2^{-(n+1)} \leq X(\omega) < (2k+2)2^{(-n+1)}, \ k = 0, 1, \ldots (n+1)2^{n+1} - 1$$

$$= n+1, \text{ if } X(\omega) \geq n+1.$$

Comparing with $X_n(\omega)$, $X_{n+1}(\omega) \geq X_n(\omega)$ for all n and ω. Since $X_n(\omega) \leq X(\omega)$ and $|X_n(\omega) - X(\omega)| < 2^{-n}$ for $X(\omega) \leq n$. as $n \to \infty$, $0 \leq X_n(\omega) \uparrow X(\omega)$ for all $\omega \in \Omega$.

Theorem:

Continuous real valued functions on R are Borel functions.

Proof:

Let $g(\cdot)$ be a continuous function from R to R. Consider $g(X)$, where X is $\mathcal{B}$-measurable function from R to R. There exists a sequence $\{X_n\}$ of elementary functions converging uniformly to X. Since X_n takes only a countable number of distinct values $g(X_n)$ also takes only a countable number of distinct values and is therefore an elementary function from R to R. Since g is continuous as $X_n(\omega) \to X(\omega)$, $g(X_n(\omega)) \to g(X(\omega))$

Thus, $g(X)$ is a limit of elementary Borel-measurable functions and is hence Borel-measurable.

Take X to be the identity function with $X(\omega) = \omega$. This is Borel function. Then $g(X)(\omega) = g(\omega)$. Since $g(X)$ is shown to be Borel measurable for any measurable function X, $g(\cdot)$ is a Borel function.

I M.Sc, MATHEMATICS

TITLE OF THE PAPER: PROBABILITY THEORY

PAPER CODE : 18KP2M08

UNIT-II

DEFINITION- PROBABILITY:

When an experiment is performed it may result in one of possibly several outcomes. With every event A consisting of one or more outcomes of an experiment. we may associate a numerical quantity. called the probability A, denoted by $P(A)$ or P_A. which will measure the 'chance' that the event A will occur.

(i) If A occur certainly. $P(A)=1$

(ii) If A will not occur. $P(A)=0$

(iii) If $P(A)=\frac{1}{2}$. there is 50% chance that A would occur.

DEFINITION- PROBABILITY FUNCTION:

Let $\mathcal{A}$ be the σ-field of events associated with the outcomes of certain experiment and Ω be the sure event. Probability is a real valued set function $P(\cdot)$ defined on $\mathcal{A}$ such that for any $A_1, A_2, \ldots \in \mathcal{A}$

(i) $P(A)\geq 0$. (Non-negative)

(ii) $P(\Omega)=1$. (Normed)

(iii) If $A=\sum_{i=1}^{n} A_i$. $P(A)=\sum_{i=1}^{n} P(A_i)$ (Finite additivity)

(iv) If $A=\sum_{i=1}^{\infty} A_i$. $P(A)=\sum_{i=1}^{\infty} P(A_i)$ (σ-Additivity)

The probability that have been assigned to various events by some method and it satisfies the above conditions (i) to (iv) are functions and such functions are called as probability function.

In the triplet $(\Omega, \mathcal{A}, P)$. Ω is the space of outcomes. $\mathcal{A}$ is the σ-field of events associated with an experiment and P is a probability function defined on $\mathcal{A}$. This triplet is called a probability space.

If a function μ defined on $(\Omega, \mathcal{A})$ possess properties (i) and (iv) it is called a measure. If $\mu(\Omega)$

is finite, μ is said to be a finite measure. Thus probability is a normed or scaled measure.

If $\mu(\Omega)=\infty$, but $\Omega=\sum A_i$ such that $\mu(A_i)<\infty$, then μ is called a σ-finite measure.

Measures such as length, defined on $(R, \mathcal{B})$ are σ-finite. The triplet $(\Omega, \mathcal{A}, \mu)$ is called the <u>measure space.</u>

Probability function P on $\mathcal{A}$ is also called probability distribution on Ω. It

An arbitrary union of events $A_1, A_2, \ldots, A_n$ can be converted into union of disjoint events $B_1, B_2, \ldots, B_n$. Hence

$P(\bigcup_{i=1}^{n} A_i) = P(\sum_{i=1}^{n} B_i) = \sum_{i=1}^{n} PB_i$ and

$P(\bigcup_{i=1}^{n} A_i) = P(A_1) + P(A_1^c A_2) + P(A_1^c A_2^c A_3) + \cdots + P(A_1^c A_2^c \ldots A_{n-1}^c A_n)$.

Now $A_1^c A_2 = (\Omega - A_1) A_2 = A_2 - A_1 A_2$

$P(A_1^c A_2) = P(A_2 - A_1 A_2) = P(A_2) - P(A_1 A_2)$.

$A_1^c A_2^c A_3 = (\Omega - A_1) A_2^c A_3 = A_2^c A_3 - A_1 A_2^c A_3 = (\Omega - A_2) A_3 - A_1 A_3 (\Omega - A_2)$

$= A_3 - A_2 A_3 - A_1 A_3 + A_1 A_2 A_3$.

$P(A_1^c A_2^c A_3) = P(A_3) - P(A_2 A_3) - P(A_1 A_3) + P(A_1 A_2 A_3)$.

$P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 A_2) - P(A_1 A_3) - P(A_2 A_3)$
$+ P(A_1 A_2 A_3)$

and hence,

$P(A_1 \cup A_2 \cup A_3 \cup A_4) = \sum_{i=1}^{4} P(A_i) - \sum_{i<j}\sum P(A_i A_j) + \sum_{i<j<k}\sum\sum P(A_i A_j A_k)$
$- P(A_1 A_2 A_3 A_4)$.

Continuing in this manner,

$P(\bigcup_{i=1}^{n} A_i) = S_1 - S_2 + S_3 - \cdots (-1)^{n-1} S_n$.

where $S_1 = \sum P(A_i)$, $S_2 = \sum_{i<j\le n}\sum P(A_i A_j)$, $S_3 = \sum_{i<j<k\le n}\sum\sum P(A_i A_j A_k)$

If $\{A_i\}$ is a partition of Ω, $B = \sum_i B A_i$ and $PB = \sum_i P(BA_i)$

Here $\{A_i\}$ may be a finite or a countable partition.

DEFINITION - DISCRETE PROBABILITY SPACE:

In a Sample space Ω, if the class events of $\mathcal{A}$ is generated by countable partition of subsets of Ω then $(\Omega, \mathcal{A}, P)$ is called the Discrete probability space.

If $\mathcal{A} = \sigma\{A_i\}$ where $\{A_i\}$ is a partition of Ω, any event A is Countable union of $\{A_i\}$. Hence the probability of A can be determined knowing the probabilities p_i of A_i. If Ω is Countable, A_i could be singletons $\{\omega_i\}$ $(i=1,2,\ldots)$. The set $\{p_1, p_2, \ldots\}$ is called the probability distribution on Ω.

FINITE PROBABILITY SPACE - BINOMIAL DISTRIBUTION:

Suppose n coins are tossed once or one coin is tossed n times. In both the cases, the space outcomes Ω can be represented by the set of all n-tuples of the form (HHTHTT....). Since each of the components may be either H or T, there are $N = 2^n$ n-tuples. If all these have the same probability, ViZ., $p_i = 2^{-n}$ $(i=1,2,\ldots,N)$. we can calculate the probabilities of events of the form.

A = (k heads and n-k tails turn up).

B = (atleast one head turns up), etc.

The number of outcomes favourable to A are $\binom{n}{k}$ and those favourable to B are all n-tuples except the point (TT...T). Hence

$$P(A) = \binom{n}{k} 2^{-n}, \quad P(B) = 1 - 2^{-n}$$

Suppose define the probability of the sample point ω_i = (HTH....) containing k, H's and (n-k), T's to be

$$P_i = p^k q^{n-k} \; (q = (1-p)) \; (i = 1, 2, \ldots, N).$$

Now, $P(B) = 1 - P(TT\ldots T) = 1 - q^n$ and $P(A) = p^k q^{n-k} \binom{n}{k}$.

Since all the $\binom{n}{k}$ points favourable to A have the same probability $p^k q^{n-k}$ The probabilities of events such as (the no. of heads exceeds the no. of tails) can be calculated.

on Ω suppose we define a r.v. X denoting the number of heads turning up; X is a mapping from Ω to $\{0, 1, 2, \dots, n\} = \Omega'$. In general

$P[X=k] = \binom{n}{k} p^k q^{n-k}$ of Ω'. The above probability. This set of probabilities on Ω' represents a distribution of probability "mass" on Ω. The range space of X called the probability distribution of X. $P[X=k]$ is the k^{th} term in the expansion of the binomial $(q+p)^n$. It is called the binomial distribution and X is called the binomial r.v.

Example - Hypergeometric Distribution:

From a lot containing n items suppose r items are chosen. This may be done in $\binom{n}{r}$ ways. Suppose there are n_1 defective items in the lot. The no. of defective items among "r" chosen may be k $(=0, 1, \dots, n_1)$.

Assuming that all the $\binom{n}{r}$ outcomes are equally likely, the number of outcomes favourable to the event 'k defectives in the samples' is $\binom{n_1}{r}\binom{n_2}{r-k}$ where $n_2 = n - n_1$ is the no. of nondefective items. This may be seen from the fact that k defective can be chosen from n_1 defectives in $\binom{n_1}{k}$ ways and the remaining $(r-k)$ non-defectives can be chosen from the n_2 non-defectives in $\binom{n_2}{r-k}$ ways.

Hence $P_k = \binom{n_1}{k}\binom{n_2}{r-k} \Big/ \binom{n}{r}$, $k = 0, 1, 2, \dots \min(n_1, r)$

$\{P_k\}$ represents a probability distribution on the range space

$\Omega' = \{0, 1, \ldots, \min(n, r)\}$ of X, representing the number of defectives observed in a sample of size r chosen from a lot of size n containing n defectives. This distribution is called the Hypergeometric distribution and X the Hypergeometric r.v.

Example- MULTINOMIAL DISTRIBUTION:

Suppose there are n cells and r balls each one of which is to be placed in one of these n cells. Total number of outcomes is $N = n^r$. If all these are equally likely, the probability $p(r_1, r_2, \ldots, r_n)$ of the event that the n cells contain $r_1, r_2, \ldots, r_n$ $(\sum r_i = r)$ balls respectively can be obtained as.

$p(r_1, r_2, \ldots, r_2, \ldots, r_n) = \frac{r!}{r_1! \ldots r_n!} \cdot \frac{1}{n^r} \longrightarrow (1)$ by noting that the no. of outcomes favorable to the event is the product of the no. of combinations of r_1 balls out of a total of r balls in the first cell, r_2 balls out of the remaining $r - r_1$ balls in the second cell and r_3 balls out of the remaining $r - r_1 - r_2$ balls in the third cell and so on.

But $\binom{r}{r_1}\binom{r - r_1}{r_2} \cdots \binom{r - r_1 - r_2 - \cdots - r_{n-2}}{r_{n-1}} = \frac{r!}{r_1! r_2! \ldots r_n!}$.

With every outcome of the experiment, we may associate an n-dimensional r.v $X = (X_1, X_2, \ldots, X_n)$, where X_i denotes the number of balls in the i-th cell $(i = 1, 2, \ldots, n)$. In the above probability space, (1) gives us the probabilities of the events $[X_1 = r_1, X_2 = r_2, \ldots, X_n = r_n] = [X = r]$. For different vectors r, such that $r \geq 0$, $\sum r_i = r$. The set of all such vectors r denotes the range space of X. By assuming probabilities $P[X = r] = \frac{r!}{r_1! \ldots r_n!} p_1^{r_1} \cdots p_n^{r_n}$, $(\sum p_i = 1)$ on this range space of X, we get a distribution called the Multinomial distribution. It can be seen that the above expression is a term in the expansion of the multinomial $(p_1 + p_2 + \cdots + p_n)^r$.

CLASSICAL OCCUPANCY PROBLEM:

Suppose r balls are to be placed in n cells so that each ball may occupy anyone of the cells. If the balls are distinguishable the number of possible arrangements is n^r. If all these are equally likely, the probability that a particular cell is empty is $(n-1)^r/n^r$. Similarly the probability that two preassigned cells are empty (null) is $(n-2)^r/n^r$. Let $P_m(r,n)$ be the probability that exactly m cells are empty. To calculate $P_m(r,n)$ we may proceed as follows:

$$P_0(r,n) = 1 - P(\text{at least one of the cells is empty}).$$
$$= 1 - P\left(\bigcup_{k=1}^{n} A_k\right)$$

where A_k is the event that the cell marked is empty.

$$P\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} PA_k - \sum_{j<k} P(A_jA_k) + \sum_{j<k<l} P(A_jA_kA_l)$$

Since $P(A_k) = [(n-1)/n]^r$, $P(A_jA_k) = [(n-2)/n]^r$, etc.,

$$P_0(r,n) = 1 - n\left(1-\frac{1}{n}\right)^r + \binom{n}{2}\left(1-\frac{2}{n}\right)^r - \cdots (-1)^k\binom{n}{k}\left(1-\frac{k}{n}\right)^r + \cdots$$

Then $n^r P_0(r,n)$ is the number of arrangements with no cell empty.

To determine $P_m(r,n)$ we note that any m of the n cells may be empty, which can happen in $\binom{n}{m}$ ways. Since each of the remaining $(n-m)$ cells is not empty, the number of such arrangements is $(n-m)^r P_0(r,n-m)$. Hence

$$P_m(r,n) = \binom{n}{m}(n-m)^r P_0(r,n-m)n^{-r}.$$

$$= \binom{n}{m}\sum_{k=0}^{n-m}(-1)^k\binom{n-m}{k}\left(1-\frac{m+k}{n}\right)^r.$$

Example:

Suppose $\Omega = \{0,1,2,\ldots\}$ and $P(\{x\}) = (1-\theta)\theta^x$, $(x=0,1,2,\ldots)$ $0<\theta<1$ → (A)

Then $P_x = P(\{x\}) \geq 0$, $\sum_{x\in\Omega} P_x = 1$, and on every subset A of Ω

$$P(A) = \sum_{x\in A} P_x$$

Thus $(\Omega, \mathcal{A}, P)$ where $\mathcal{A}$ is the class of all subsets of Ω, is a probability space. This distribution is called a geometric distribution with (A) as its probability mass function (p.m.f)

Example:

once again let $\Omega = \{0,1,2,\ldots\}$ and $\mathcal{A}$ be the class of all subsets of Ω. If P specifies that $P(\{x\}) = \frac{e^{-\lambda}\lambda^x}{x!}$ $(\lambda > 0)$, $(x = 0,1,2,\ldots)$ specifies a Poisson Distribution.

Soln:

Let X be a mapping from Ω to $\{0,1,2,\ldots\}$ such that

$$X^{-1}\{x\} = \{\omega : X(\omega = x)\ X(\omega) = x\} A_x.$$

Suppose $\{A_x\}$ which is partition of Ω generates $\mathcal{A}$. Let $P[X = x] = P(A_x) = P_x$. Say. Then the probability distribution on $\mathcal{A}$ is specified by the probabilities of the events $[X = x]$.

If $P[X = x] = (e^{-\lambda}\lambda^x)/x!$ then X is said to be a Poisson r.v.

If $P[X = x] = (1-\theta)\theta^x$ then X is said to be geometric r.v.

If $P[X = x] = \binom{n+x-1}{n-1} p^n q^x$, $(0 < p < 1, q = 1-p)$.

then X is called a negative binomial r.v and the corresponding probability distribution is called the negative binomial distribution.

To find the probability of the event, $[X = \text{odd number}]$

$P(X = \text{odd}) = \sum_{x=\text{odd}} P_x$ and this probability equals $e^{-\lambda}\sum_{n=0}^{\infty}\frac{\lambda^{2n+1}}{(2n+1)!}$ for a poisson random variable. Similarly

$(1-\theta)\sum_{n=0}^{\infty}\theta^{2n+1} = \frac{\theta}{1+\theta}$ for a geometric random variable.

Similarly the probabilities of the events $[X \geq x]$, $[X > x]$, $[X = \text{even}]$. when the probability distribution is specified by $P(\{x\}) = \frac{e^{-\lambda}\lambda^x}{x!}$ $(\lambda > 0)$, $(x = 0,1,2,\ldots)$, and.

$$P[X = x] = \binom{n+x-1}{n-1} p^n q^x \ (0 < p < 1, q = 1-p)$$

DEFINITION - DISCRETE RANDOM VARIABLE:

A random variable (or vector) X taking only a countable number (finite or infinite) of distinct values $(x_0, x_1, x_2, \ldots)$ say is called a Discrete random variable.

CONTINUITY PROPERTY:

LEMMA:

In the presence of

Finite Additivity:

If $A = \sum_{i=1}^{n} A_i$. $P(A) = \sum_{i=1}^{n} P(A_i)$ and $\longrightarrow$ (iii)

σ-Additivity:

If $A = \sum_{i=1}^{\infty} A_i$, $P(A) = \sum_{i=1}^{\infty} P(A_i)$ $\longrightarrow$ (iv) is equivalent to the following:

(iv)': $B_n \in \mathcal{A}$, $B_n \uparrow B \Rightarrow P(B_n) \uparrow P(B) \uparrow P(B)$ as $n \longrightarrow \infty$.

Proof:

Let $A_n = B_n - B_{n-1}$ so that A_n's are disjoint and $B_n = \sum_{i=1}^{n} A_i$. Then $B_n \uparrow \sum_{i=1}^{\infty} A_i = B$ as $n \longrightarrow \infty$.

But $P(B_n) = \sum_{i=1}^{n} P(A_i) \uparrow \sum_{i=1}^{\infty} P(A_i)$.

If (iv)' holds. $B_n \uparrow B \Rightarrow P(B_n) \longrightarrow P(B)$ and hence taking $B_n = \sum_{i=1}^{n} A_i$, $\sum_{i=1}^{\infty} P(A_i) = P\left(\sum_{i=1}^{\infty} A_i\right)$ and (iv) follows

Conversely let (iv) holds. Since

$$P(B) = \sum_{i=1}^{\infty} P(A_i) = \lim_{n} \sum_{i=1}^{n} P(A_i) = \lim_{n} P(B_n).$$

and $P(B_n) \uparrow P(B)$ as $n \longrightarrow \infty$.

Corollary:

If $A_n \in \mathcal{A}$, $A_n \uparrow A$, then $P(A_n) \downarrow P(A)$.

The corollary follows because

$A_n \downarrow A \Rightarrow A_n^c \uparrow A^c \Rightarrow P(A_n^c) \longrightarrow P(A^c) \Rightarrow \{1 - P(A_n)\} \longrightarrow \{1 - P(A)\}$
$\longrightarrow P(A_n) \to P(A)$

Corollary:

If $A_n \longrightarrow A$, then $P(A_n) \longrightarrow P(A)$.

Proof: $\bigcap_{k \geq n} A_k \subset A_n \subset \bigcup_{k \geq n} A_k \longrightarrow P\left(\bigcap_{k \geq n} A_k\right) \leq P(A_n) \leq P\left(\bigcup_{k \geq n} A_k\right)$.

But $\bigcap_{k \geq n} A_k \uparrow \underline{\lim} A_n$, $\bigcup_{k \geq n} A_k \downarrow \overline{\lim} A_n$.

Since $A_n \to A$, $\underline{\lim} A_n = \overline{\lim} A_n = A$.

But $P(\bigcap_{k \geq n} A_k) \uparrow P(A)$ and $P(\bigcup_{k \geq n} A_k) \downarrow P(A)$.

Hence $P(A_n) \to P(A)$.

THEOREM:

Probability function defined on all intervals of the form (a, b), $a, b \in R$, $a < b$ defines uniquely an extension to the minimal field containing all the intervals.

Proof:

Let $\mathcal{D}_1$ be the class of all intervals of the form $(a, b]$ and $\mathcal{D}_0$ be the minimal field containing $\mathcal{D}_1$. It contains all finite sums of sets of $\mathcal{D}_1$ and complements of sets of $\mathcal{D}_1$ (ie) sets of the form $(-\infty, a] \cup (b, \infty)$.

Let $A \in \mathcal{D}_0$ and $A = \sum_{k=1}^{m} A_k$, $A_k \in \mathcal{D}_1$. Then $P(A) = \sum_k PA_k$.

If A is also equal to $\sum_{j=1}^{n} A_j'$, $A_j' \in \mathcal{D}_1$, then $P(A) = \sum_j PA_j'$.

We have to prove that, $\sum_k PA_k = \sum_j PA_j'$.

Since P on $\mathcal{D}_1$ is additive and

$$A_j' = AA_j' = \sum_k A_k A_j'$$

$$A_k = AA_k = \sum_j A_k A_j'$$

It follows that,

$$\sum_j P(A_j') = \sum_j P\left(\sum_k (A_k A_j')\right) = \sum_j \sum_k P(A_k A_j')$$

$$= \sum_{j,k} P(A_j' A_k) = \sum_k P\left(\sum_j (A_j' A_k)\right) = \sum_k P(A_k).$$

Since $P(A^c) = 1 - PA$, for all $A \in \mathcal{D}_1$.

DISTRIBUTION OF BOREL FUNCTIONS OF RANDOM VARIABLES:

If g is any Borel function of X,

$$P[g(X(\omega)) \in B] = P[X(\omega) \in g^{-1}(B)] = P_X(g^{-1}(B))$$

Since g is a Borel function $g^{-1}(B)$ is a Borel set. Hence the probability distribution of X determines uniquely the probability distribution of any Borel function of X.

Taking $g(X) = 2X$, in the binomial expression,

$$P[2X = k] = P[X = k/2] = 0, \text{ if } k \text{ is odd.}$$

$$\binom{n}{k/2} p^{k/2} q^{n-k/2}, \text{ if } k \text{ is even.}$$

Now $P[X^2 \le k] = P[-\sqrt{k} \le X \le +\sqrt{k}]$.

$$P[a \le X^2 \le b] = P[(-\sqrt{b} \le x \le -\sqrt{a}) \cup (+\sqrt{a} \le X \le +\sqrt{b})]$$

hence we can determine the distribution of X^2 knowing the distribution of X. Similarly we can determine the distributions of e^x, $\log X$ and other one-valued functions $Y = g(X)$. In fact if g is increasing

$$P[g(X) \le y] = P[X \le g^{-1}(y)].$$

If $f(x)$ is the probability density function (p.d.f) of X and if g^{-1} exists, g, g^{-1} are differentiable, using the calculus of transformation of variables.

$$P[g(x_1) \le g(X) \le g(x_2)] = \int_{x_1}^{x_2} f(x)\,dx = \int_{y_1}^{y_2} f(h(y))\,h'(y)\,dy.$$

where $y_1 = g(x_1)$, $y_2 = g(x_2)$, $g^{-1}(y) = x = h(y)$.