MAT2001
Statistics for Engineers

Module 2
Random Variables

Syllabus

Random variables:
Introduction -random variables-Probability mass Function, distribution and density functions - joint Probability distribution and joint density functions-Marginal, conditional distribution and density functions- Mathematical expectation, and its properties Covariance , moment generating function – characteristic function.

Probability Theory

Probability theory had its origin in the analysis of certain games of chance that were popular in the seventeenth century. It has since found applications in many branches of Science and Engineering and this extensive application makes it an important branch of study. Probability theory, as a matter of fact, is a study of random or unpredictable experiments and is helpful in investigating the important features of these random experiments.

- **Random Experiment**
- **Sample Space**
- **Events**

Random Experiment, Sample Space and Event

<u>Problem</u> (Non-deterministic)

<u>Random Experiment</u>

$S \to$ Sample Space (All possible Outcomes)

$\Omega = \mathcal{P}(S) =$ Power set of S

$= \{A : A \subseteq S\}$

<u>Event</u>: $\to$ a subset of a sample space S

A is called an event, if $A \subseteq S$

(i.e) $A \in \Omega = \mathcal{P}(S)$

Mathematical Definition of Probability

Let E - Experiment with Sample Space S.

and let A - an event. Here S is finite

n(S) is finite

$$\boxed{P(A) = \frac{n(A)}{n(S)} = \frac{|A|}{|S|}}$$

where,

$\left.\begin{matrix} n(A) \\ |A| \end{matrix}\right\}$ = no/. of elements in A

$$P(A) = \frac{n(A)}{n(S)} = \frac{\text{Number of cases favourable to } A}{\text{Exhaustive number of cases in } S}$$

Example:

1. $S = \{H, T\} \Rightarrow |S| = 2$

$A = \{H\} \Rightarrow |A| = 1$

$P(A) = 1/2$

2. $S = \{1, 2, \ldots, 6\} \Rightarrow |S| = 6$

$O = \{1, 3, 5\} \Rightarrow |O| = 3$

$P(O) = 1/2$

Statistical Definition of Probability

n – Trials

A = Expected Event

n_A = No. of occurrence of A

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

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

Axiomatic Definition of Probability

Probability | Probability function |

Probability Measure

Let E - Exp with S.S 'S'

A function

Let $A \subseteq S$

$$P : \mathcal{P}(S) \longrightarrow [0,1]$$

is called probability, if P satisfies

(i). $P(S) = 1$ [(or) $P(\Phi) = 0$]

(ii). Addition theorem of Probability.

Mutually Exculsive Events

Let A & B are 2 events

A and B are said to be mutually exculsive events, if $A \cap B = \Phi$

(or) $P(A \cap B) = P(\Phi) = 0$

Addition Theorem of Probability

Let A & B be an events

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

Suppose A & B are mutually exclusive

then

$$\boxed{P(A \cup B) = P(A) + P(B)}$$

In General

Suppose $A_1, A_2, \ldots, A_n, \ldots$ are pair-wise mutually exclusive events,

then $P(A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots) = P(A_1) + P(A_2) + \cdots + \cdots + P(A_n) + \cdots$

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

Sure Event

Sample Event

$P(S) = 1$

$\phi \rightarrow$ Impossible Event

Theorem 1

The probability of the impossible event is zero, i.e., if ϕ is the subset (event) containing no sample point, $P(\phi) = 0$.

Theorem 2

If $\overline{A}$ is the complementary event of A, $P(\overline{A}) = 1 - P(A) \le 1$.

Theorem 3

If A and B are any 2 events, $P(A \cup B) = P(A) + P(B) - P(A \cap B) \le P(A) + P(B)$

Theorem 4

If $B \subset A$, $P(B) \le P(A)$.

Conditional Probability

The conditional probability of an event B, assuming that the event A has happened, is denoted by $P(B/A)$ and defined as

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

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

$$\boxed{P(A \cap B) = P(A) \times P(B/A)}$$

Conditional Probability

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

<u>Example</u>

$S = \{1, 2, \ldots, 6\}$

$E = \{2, 4, 6\} \Rightarrow P(E) = 1/2$

$O = \{1, 3, 5\} \Rightarrow P(O) = 1/2$

$A = \{1\}$ $\quad A \cap O = \{1\} \Rightarrow P(A \cap O) = 1/6$

$P(A) = 1/6$

$$P(A/O) = \frac{P(A \cap O)}{P(O)} = \frac{1/6}{1/2} = 1/3$$

$$P(A/E) = \frac{P(A \cap E)}{P(E)} = \frac{P(\phi)}{P(E)} = 0$$

Product theorem of probability

$$P(A \cap B) = P(A) \times P(B/A)$$

The product theorem can be extended to 3 events *A*, *B* and *C* as follows:

$$P(A \cap B \cap C) = P(A) \times P(B/A) \times P(C/A \text{ and } B)$$

Independent Events

$$P(A \cap B) = P(A) \times P(B/A)$$

If A and B are independent to eachother,

$$P(B/A) = P(B)$$

$$\boxed{P(A \cap B) = P(A) \times P(B)}$$

Independent Events

If $P(A \cap B) = P(A) \times P(B)$,

the events A and B are said to be independent (pairwise independent).

In General,

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \times P(A_2) \times \ldots \times P(A_n)$$

Theorem 1

If the events A and B are independent, the events $\overline{A}$ and B (and similarly A and $\overline{B}$) are also independent.

Theorem 2

If the events A and B are independent, then so are $\overline{A}$ and $\overline{B}$.

Random Variables (X)

$X(s) = \underline{1}$

$R_X = \{1\}$

$$X : S \longrightarrow \mathbb{R}$$

$$\{X \le x\} = \{s \in S : \underline{X(s) \le x}\}$$

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

Example:

$S = \{1, 2, 3, \ldots, 6\}$

$A = \{1, 2, 3\}$, $B = \{4, 5, 6\}$

$P(A)$ & $P(B) = ?$

$$X : S \to \mathbb{R} \implies R_X = \{-3, 3\}$$

$$X(s) = \begin{cases} -3 ; \text{ if } s \in \{1, 2, 3\} \\ +3 ; \text{ if } s \in \{4, 5, 6\} \end{cases}$$

$$\{X = -3\} = \{s \in S : X(s) = -3\}$$
$$= \{1, 2, 3\}$$
$$= A$$

$P(A) = P(X = -3)$

$P(B) = P(X = 3)$

Random Variables

Definition: A random variable (abbreviatively RV) is a function that assigns a real number $X(s)$ to every element $s \in S$, where S is the sample space corresponding to a random experiment E.)

Hereafter, R_x will be referred to as ***Range space***.

Similarly $\{X \le x\}$ represents the subset $\{s: X(s) \le x\}$ and hence an event associated with the experiment.

Discrete Random Variable

Continuous Random Variable

$X = x_i$, where $x_i \in \{x_1, x_2, \ldots, x_n\}$

$X = x$, where $x \in \underbrace{[a, b]}_{= R_x} \ni a, b \in \mathbb{R}$

Discrete Random Variable	Continuous Random Variable

Discrete Random Variable

X – DRV

$X = x_1, x_2, \ldots, x_n$

$x_i \in \mathbb{R}$

Probability Mass function (p.m.f)

x_i :	x_1	$\ldots$	x_n
p_i :	p_1	$\ldots$	p_n

(i) p_i's ≥ 0 (+ve ness)

(ii). $\sum_{i=1}^{n} p_i = \underline{1}$

p_i's are p.m.f

$\{x_i, p_i\}_{i=1}^{n}$ is called probl. distribution of a DRV X

Continuous Random Variable

X – CRV

$X = x \in [a, b] = R_X$

$[a, b] \subseteq \mathbb{R}$

Prob. Density function (p.d.f)

$f(x)$ is said p.d.f if (i). $f(x) \geq 0$, $\forall x \in R_X$ (+ve ness)

(ii) $\displaystyle\int_{R_X} f(x)\, dx = \underline{1}$

$\{x, f(x)\}_{x \in R_X}$ is called P.D of a CRV X

Discrete Random Variable

If X is a random variable (RV) which can take a finite number or countably infinite number of values, X is called a discrete RV.

Probability Function

If X is a discrete RV which can take the values $x_1, x_2, x_3, \ldots$ such that $P(X = x_i) = p_i$, then p_i is called the *probability function or probability mass function or point probability function*, provided p_i $(i = 1, 2, 3, \ldots)$ satisfy the following conditions:

(i) $p_i \geq 0$, for all i, and

(ii) $\sum_i p_i = 1$

The collection of pairs $\{x_i, p_i\}$, $i = 1, 2, 3, \ldots$, is called *the probability distribution of the* RV X, which is sometimes displayed in the form of a table as given below:

$X = x_i$	$P(X = x_i)$
x_1	p_1
x_2	p_2
$\vdots$	$\vdots$
x_r	p_r
$\vdots$	$\vdots$

Continuous Random Variable

If X is an RV which can take all values (i.e., *infinite number* of values) in an interval, then X is called a *continuous* RV.

Probability Density Function

If X is a continuous RV such that

$$P\left\{x-\frac{1}{2}\mathrm{d}x \le X \le x+\frac{1}{2}\mathrm{d}x\right\} = f(x)\mathrm{d}x$$

then $f(x)$ is called the *probability density function* (shortly denoted as pdf) of X, provided $f(x)$ satisfies the following conditions:

(i) $f(x) \ge 0$, for all $x \in R_x$, and

(ii) $\int_{R_X} f(x)\mathrm{d}x = 1$

Moreover, $P(a \le X \le b)$ or $P(a < X < b)$ is defined as

$$P(a \le X \le b) = \int_a^b f(x)\mathrm{d}x.$$

When X is a continuous RV

$$P(X = a) = P(a \le X \le a) = \int_a^a f(x)\,\mathrm{d}x = 0$$

Cumulative Distribution Function (cdf)

If X is an R V, discrete or continuous, then $P(X \leq x)$ is called the *cumulative distribution function* of X or *distribution function* of X and denoted as $F(x)$.
If X is discrete,

$$F(x) = \sum_{\substack{j \\ x_j \leq x}} p_j$$

If X is continuous,

$$F(x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(x)\,\mathrm{d}x$$

Properties of the *cdf* $F(x)$

1. $F(x)$ is a non-decreasing function of x, i.e., if $x_1 < x_2$, then $F(x_1) \leq F(x_2)$.
2. $F(-\infty) = 0$ and $F(\infty) = 1$.
3. If X is a discrete R V taking values $x_1, x_2, \ldots$, where $x_1 < x_2 < x_3 < \ldots < x_{i-1} < x_i < \ldots$, then $P(X = x_i) = F(x_i) - F(x_{i-1})$.
4. If X is a continuous R V, then $\dfrac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x)$, at all points where $F(x)$ is differentiable.

Special Distributions

Discrete Distributions

1. *Binomial Distribution*
2. *Poisson Distribution*

Continuous Distributions

1. *Normal Distribution*
2. *Exponential Distribution*
3. *Gamma Distribution*
4. *Weibull Distribution*

X - RV

$$\underset{\downarrow \text{Mean}}{E(X)} = \begin{cases} \dfrac{\sum_{i=1}^{n} x_i p_i}{\sum p_i} = \sum_{i=1}^{n} x_i p_i & ;\text{if } X \text{ is DRV} \\ \int_{-\infty}^{\infty} x f(x)\, dx & ;\text{if } X \text{ CRV} \end{cases}$$

$$E(X^2) = \begin{cases} \sum_{i=1}^{n} x_i^2 p_i \; ; & \text{if } X - \text{DRV} \\ \int_{-\infty}^{\infty} x^2 f(x)\, dx ; & \text{if } X - \text{CRV} \end{cases}$$

Let '$g(x)$' be a function on a RV X.

$$\boxed{E(g(x)) = \begin{cases} \sum_{i=1}^{n} g(x_i) \cdot p_i \; ; & \text{if } X - \text{DRV} \\ \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx \; ; & \text{if } X - \text{CRV} \end{cases}}$$

Suppose $g(x) = x$ (Identity function)

$E(g(x)) = E(X)$

Suppose $g(x) = x^2$

then $E(g(x)) = E(X^2)$

<u>Properties:</u>

$E(aX) = a\,E(X)$

$E(aX + bY) = a\,E(X) + b\,E(Y)$

$E(X+Y) = E(X) + E(Y)$

$E(1) = \underline{1}$

$E(a) = a$

$$\text{Variance} = \frac{1}{N}\sum x_i^2 p_i - \left(\frac{1}{N}\sum x_i p_i\right)^2 = E(X^2) - (E(X))^2$$

$$= \frac{1}{N}\sum (x_i - \bar{x})^2 p_i = \sum (x_i - E(x))^2 p_i$$

$$= E(X - E(X))^2$$

$$\boxed{\begin{aligned} Var(X) &= E(X^2) - (E(X))^2 \\ \sigma_X^2 &= E\left[(X - E(X))^2\right] \end{aligned}}$$

$$S.D(X) = \sigma_X = +\sqrt{Var(X)}$$

Mathematical Expectation of One Dimensional Random Vairable

Definitions: If X is a discrete RV, then ***the expected value*** or the mean value of $g(X)$ is defined as

$$E\{g(X)\} = \sum_i g(x_i)p_i,$$

where $p_i = P(X = x_i)$ is the probability mass function of X.

If X is a continuous RV with pdf $f(x)$, then

$$E\{g(X)\} = \int_{R_X} g(x)f(x)\,dx$$

Mean, Variance and Standard Deviation of a Random Variable

Two expected values which are most commonly used for characterising a RV X are ***its mean*** μ_X and ***variance*** σ^2_X, which are defined as follows:

$$\mu_X = E(X)$$
$$= \sum_i x_i \, p_i \text{, if } X \text{ is discrete}$$
$$= \int_{R_X} x f(x)\, dx \text{, if } X \text{ is continuous}$$

$$\mathrm{Var}(X) = \sigma^2_X = E\{(X - \mu_X)^2\}$$
$$= \sum_i (x_i - \mu_X)^2 \, p_i \text{, if } X \text{ is } \textit{discrete}$$
$$\int_{R_x} (x - \mu_X)^2 f(x)\, dx \text{, if } X \textit{ is continuous}$$

$$Var(X) = E(X^2) - \{E(X)\}^2$$

The square root of variance is called ***the standard deviation.***

Properties of Expected Value:

1). $E(aX) = a\,E(X)$

2). $Var(aX) = a^2\,Var(X)$

3). $E(X+Y) = E(X) + E(Y)$

4). If X and Y are independent RVs, then $E(XY) = E(X) \cdot E(Y)$

Example:

If the random variable X takes the values 1, 2, 3 and 4 such that $2P(X = 1) = 3P(X = 2) = P(X = 3) = 5P(X = 4)$, find the probability distribution and cumulative distribution function of X.

Solution:

Let $P(X = 3) = 30K$. Since $2P(X = 1) = 30K$, $P(X = 1) = 15K$.
Similarly $P(X = 2) = 10K$ and $P(X = 4) = 6K$.
Since $\Sigma\, p_i = 1$, $15K + 10K + 30K + 6K = 1$.

$$\therefore\ K = \frac{1}{61}$$

The probability distribution of X is given in the following table:

$X = i$	1	2	3	4
p_i	$\frac{15}{61}$	$\frac{10}{61}$	$\frac{30}{61}$	$\frac{6}{61}$

The cdf $F(x)$ is defined as $F(x) = P(X \le x)$. Accordingly the cdf for the above distribution is found out as follows:

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

When $1 \le x < 2$, $F(x) = P(X = 1) = \frac{15}{61}$

When $2 \le x < 3$, $F(x) = P(X = 1) + P(X = 2) = \frac{25}{61}$

When $3 \le x < 4$, $F(x) = P(X = 1) + P(X = 2) + P(X = 3) = \frac{55}{61}$

When $x \ge 4$, $F(x) = P(x = 1) + P(x = 2) + P(x = 3) + P(x = 4) = 1$.

Exercise:

A random variable X has the following probability distribution.

x:	-2	-1	0	1	2	3
$p(x)$:	0.1	K	0.2	$2K$	0.3	$3K$

(a) Find K, (b) Evaluate $P(X < 2)$ and $P(-2 < X < 2)$, (c) find the cdf of X and

(d) evaluate the mean of X.

Exercise:

A random variable X has the following probability distribution.

x :	0	1	2	3	4	5	6	7
$p(x)$:	0	K	$2K$	$2K$	$3K$	K^2	$2K^2$	$7K^2 + K$

Find (i) the value of K, (ii) $P(1.5 < X < 4.5/X > 2)$

Example:

If $p(x) = \begin{cases} x\,e^{-x^2/2} & x \ge 0 \\ 0 & x < 0 \end{cases}$

(a) show that $p(x)$ is a pdf (of a continuous RV X.)
(b) find its distribution function $P(x)$.

a)

(i). $I = \int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} \underline{f(x)}\,dx + \int_{0}^{\infty} f(x)\,dx$

$\quad \underline{f(x)} \to 0$

$I = \int_0^\infty x \cdot e^{-x^2/2}\,dx$

Let $t = x^2 \Rightarrow x = \sqrt{t}$ | $x = 0 \Rightarrow t = 0$

$dt = 2x\,dx$ | $x = \infty \Rightarrow t = \infty$

$I = \frac{1}{2}\int_0^\infty e^{-t/2}\,dt$

$I = \cancel{\tfrac{1}{2}}\left[\frac{e^{-t/2}}{-\cancel{\tfrac{1}{2}}}\right]_0^\infty$

$= -\left[0 - 1\right]$

$= 1$

b).

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

$-\infty < x < 0, \quad F(x) = 0$

$0 \le x < \infty, \quad F(x) = \int_0^x f(x)\,dx$

$= \int_0^x x \cdot e^{-x^2/2}\,dx$

$F(x) = 1 - e^{-x^2/2} \qquad \boxed{t = x^2}$

$F(x) = \begin{cases} 0; & -\infty < x < 0 \\ 1 - e^{-x^2/2}; & 0 \le x < \infty \end{cases}$

Solution:

(a) If $p(x)$ is to be a pdf, $p(x) \geq 0$ and

$$\int_{R_X} p(x)\,dx = 1$$

Obviously, $p(x) = xe^{-x^2/2} \geq 0$, when $x \geq 0$

Now $\int_0^\infty p(x)dx = \int_0^\infty xe^{-x^2/2}\,dx = \int_0^\infty e^{-t}\,dt$ (putting $t = x^2/2$)

$$= 1$$

$\therefore$ $p(x)$ is a legitimate pdf of a RV X.

$$F(x) = P(X \leq x) = \int_0^x f(x)dx$$

$\therefore$ $F(x) = 0$, when $x < 0$

and $F(x) = \int_0^X xe^{-x^2/2}\,dx = 1 - e^{-x^2/2}$, when $x \geq 0$.

Example:

If the density function of a continuous RV X is given by

$$
\begin{aligned}
f(x) &= ax, && 0 \le x \le 1 \\
&= a, && 1 \le x \le 2 \\
&= 3a - ax, && 2 \le x \le 3 \\
&= 0, && \text{elsewhere}
\end{aligned}
$$

(i) find the value of a
(ii) find the cdf of X

Solution:

(i) Since $f(x)$ is a pdf, $\int_{R_X} f(x)\mathrm{d}x = 1$

i.e., $\int_0^3 f(x)\mathrm{d}x = 1$

i.e., $\int_0^1 ax\mathrm{d}x + \int_1^2 a\mathrm{d}x + \int_2^3 (3a - ax)\mathrm{d}x = 1$

i.e., $2a = 1$

$\therefore$ $a = \frac{1}{2}$

(ii) $F(x) = P(X \le x) = 0$, when $x < 0$

$$F(x) = \int_0^x \frac{x}{2}\mathrm{d}x = \frac{x^4}{4}, \text{ when } 0 \le x \le 1$$

$$= \int_0^1 \frac{x}{2}\mathrm{d}x + \int_1^x \frac{1}{2}\mathrm{d}x = \frac{x}{2} - \frac{1}{4} \text{ when } 1 \le x \le 2$$

$$= \int_0^1 \frac{x}{2}\mathrm{d}x + \int_1^2 \frac{1}{2}\mathrm{d}x + \int_2^x \left(\frac{3}{2} - \frac{x}{2}\right)\mathrm{d}x = \frac{3}{2}x - \frac{x^2}{4} - \frac{5}{4}, \text{ when } 2 \le x \le 3$$

$= 1$, when $x > 3$

Exercise:

A continuous RV X that can assume any value between $x = 2$ and $x = 5$ has a density function given by $f(x) = k(1 + x)$. Find $P(X < 4)$.

Exercise:

A continuous RV X has a pdf $f(x) = kx^2e^{-x}$; $x > 0$. Find k, mean and variance.

Exercise:

A continuous RV has a pdf $f(x) = 3x^2, 0 \le x \le 1$. Find a and b such that

(i) $P(X \le a) = P(X > a)$ and

(ii) $P(X > b) = 0.05$

Two-Dimensional Random Variables

Definitions: Let S be the sample space associated with a random experiment E Let $X = X(s)$ and $Y = Y(s)$ be two functions each assigning a real number to eacl outcomes $s \in S$. Then (X, Y) is called a *two-dimensional random variable.*

If the possible values of (X, Y) are finite or countably infinite, (X, Y) is called *a two-dimensional discrete* RV. When (X, Y) is a two-dimensional discrete RV the possible values of (X, Y) may be represented as (x_i, y_j), $i = 1, 2, \ldots, m, \ldots; j = 1, 2, \ldots, n, \ldots$.

If (X, Y) can assume all values in a specified region R in the xy-plane, (X, Y) is called a *two-dimensional continuous* RV.

Probability Function of (X, Y)

If (X, Y) is a two-dimensional discrete RV such that $P(x = x_i, y = y_j) = p_{ij}$, then p_{ij} is called the *probability mass function* or simply the *probability function* of (X, Y) provided the following conditions are satisfied.

(i) $p_{ij} \geq 0$, for all i and j

(ii) $\sum_j \sum_i p_{ij} = 1$

The set of triples $\{x_i, y_j, p_{ij}\}$, $i = 1, 2, \ldots, m, \ldots, j = 1, 2, \ldots, n, \ldots$, is called *the joint probability distribution of (X, Y)*.

Joint Probability Density Function

If (X, Y) is a two-dimensional continuous RV such that.

$P\left\{x - \frac{dx}{2} \leq X \leq x + \frac{dx}{2} \text{ and } y - \frac{dy}{2} \leq Y \leq y + \frac{dy}{2}\right\} = f(x, y)\, dx\, dy$, then $f(x, y)$ is called *the joint pdf* of (X, Y), provided $f(x, y)$ satisfies the following conditions.

(i) $f(x, y) \geq 0$, for all $(x, y) \in R$, where R is the range space.

(ii) $\iint_R f(x, y)\, dx\, dy = 1$.

Moreover if D is a subspace of the range space R, $P\{(X, Y) \in D\}$ is defined as

$$P\{(X, Y) \in D\} = \iint_D f(x, y) dx\, dy. \text{ In particular}$$

$$P\{a \leq X \leq b, c \leq Y \leq d\} = \int_c^d \int_a^b f(x, y) dx\, dy$$

Cumulative Distribution Function

If (X, Y) is a two-dimensional RV (discrete or continuous), then $F(x, y) = P\{X \le x \text{ and } Y \le y\}$ is called *the cdf of* (X, Y).

In the discrete case,

$$F(x, y) = \sum_{\substack{j \\ y_j \le y}} \sum_{\substack{i \\ x_i \le x}} p_{ij}$$

In the continuous case,

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

Properties of $F(x, y)$

(i) $F(-\infty, y) = 0 = F(x, -\infty)$ and $F(\infty, \infty) = 1$

(ii) $P\{a < X < b, Y \le y\} = F(b, y) - F(a, y)$

(iii) $P\{X \le x, c < Y < d\} = F(x, d) - F(x, c)$

(iv) $P\{a < X < b, c < Y < d\} = F(b, d) - F(a, d) - F(b, c) + F(a, c)$

(v) At points of continuity of $f(x, y)$

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

Marginal Probability Distribution

$$P(X = x_i) = P\{(X = x_i \text{ and } Y = y_1) \text{ or } (X = x_i \text{ and } Y = y_2) \text{ or etc.}\}$$

$$= p_{i1} + p_{i2} + \dots = \sum_j p_{ij}$$

$P(X = x_i) = \sum_j p_{ij}$ is called *the marginal probability function of X*. It is defined for $X = x_1, x_2, \dots$ and denoted as P_{i*}. The collection of pairs $\{x_i, p_{i*}\}$, $i = 1, 2, 3, \dots$ is called the *marginal probability distribution of X*.

Similarly the collection of pairs $\{y_j, p_{*j}\}$, $j = 1, 2, 3, \dots$ is called *the marginal probability distribution of Y*, where $p_{*j} = \sum_i P_{ij} = P(Y = y_j)$.

In the continuous case,

$$P\{x - \frac{1}{2}\text{d}x \le X \le x + \frac{1}{2}\text{d}x, -\infty < Y < \infty\}$$

$$= \int_{-\infty}^{\infty} \int_{x-\frac{1}{2}\text{d}x}^{x+\frac{1}{2}\text{d}x} f(x, y)\, \text{d}x\, \text{d}y$$

$$= \left[\int_{-\infty}^{\infty} f(x, y)\, \text{d}y\right] \text{d}x$$ [since $f(x, y)$ may be treated a constant in

$(x - 1/2\text{d}x, x + 1/2\,\text{d}x)$]

$= f_X(x)\text{d}x$, say

$f_X(x) = \int_{-\infty}^{\infty} f(x, y)\text{d}y$ is called *the marginal density of X*.

Similarly, $f_Y(y) = \int_{-\infty}^{\infty} f(x, y)\,\text{d}x$ is called *the marginal density of Y*.

Note $P(a \le X \le b) = P(a \le X \le b, -\infty < Y < \infty)$

$$= \int_{-\infty}^{\infty}\int_{a}^{b} f(x, y)\,\mathrm{d}x\,\mathrm{d}y$$

$$= \int_{a}^{b}\left[\int_{-\infty}^{\infty} f(x, y)\,\mathrm{d}y\right]\mathrm{d}x = \int_{a}^{b} f_X(x)\,\mathrm{d}x$$

Similarly, $P(c \le Y \le d) = \int_{c}^{d} f_Y(y)\,\mathrm{d}y$

Conditional Probability Distribution

$P\{X = x_i/Y = y_j\} = \dfrac{P\{X = x_i, Y = y_j\}}{P\{Y = y_j\}} = \dfrac{p_{ij}}{p_{\bullet j}}$ is called *the conditional probability function of* X, given that $Y = y_j$.

The collection of pairs, $\left\{x_i, \dfrac{p_{ij}}{p_{\bullet j}}\right\}$ $i = 1, 2, 3, \ldots,$

is called *the conditional probability distribution of X, given* $Y = y_j$.

Similarly, the collection of pairs, $\left\{Y_j, \dfrac{p_{ij}}{p_{i\bullet}}\right\}$, j = 1, 2, 3, …, is called the *conditional probability distribution of Y given* $X = x_i$. In the continuous case,

$$P\left\{x - \frac{1}{2}\,dx \le X < x + \frac{1}{2}\,dx / Y = y\right\}$$

$$= P\left\{x - \frac{1}{2}\,dx \le X \le x + \frac{1}{2}\,dx / y - \frac{1}{2}\,dy \le Y \le y + \frac{1}{2}\,dy\right\}$$

$$= \frac{f(x, y)\,dx\,dy}{f_Y(y)\,dy} = \left\{\frac{f(x, y)}{f_Y(y)}\right\} dx.$$

$\dfrac{f(x, y)}{f_Y(y)}$ is called *the conditional density of X, given Y*, and is denoted by $f(x/y)$.

Similarly, $\dfrac{f(x, y)}{f_X(x)}$ is called *the conditional density of Y*, given X, and is denoted by $f(y/x)$.

Independent RVs

If (X, Y) is a two-dimensional discrete RV such that $P\{X = x_i / Y = y_j\} = P(X = x_i)$

i.e., $\frac{p_{ij}}{p_{\bullet j}} = p_{i\bullet}$, i.e., $p_{ij} = p_{i\bullet} \times p_{\bullet j}$ for all i, j then X and Y are said to be independent RVs.

Similarly if (X, Y) is a two-dimensional continuous RV such that $f(x, y) = f_X(x) \times f_Y(y)$, then X and Y are said to be independent RVs.

Random Vectors

Definitions: A vector X: $[X_1, X_2, \ldots, X_n]$ whose components X_i are RVs is called a *random vector*. $(X_1, X_2, \ldots, X_n)$ can assume all values in some region R_n of the n-dimensional space. R_n is called the *range space*.

Mathematical Expectation of Two Dimensional Random Vairable

Expected Values of a Two-Dimensional RV

If (X, Y) is a two-dimensional discrete RV with joint probability mass function p_{ij}, then $E\{g(X, Y)\} = \sum_j \sum_i g(x_i, y_i\} p_{ij}$.

If (X, Y) is a two-dimensional continuous RV with joint pdf $f(x, y)$, then

$$E\{g(X, Y)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y)\, dx\, dy$$

Example:

Three balls are drawn at random without replacement from a box containing 2 white, 3 red and 4 black balls. If X denotes the number of white balls drawn and Y denotes the number of red balls drawn, find the joint probability distribution of (X, Y).

Solution:

As there are only 2 white balls in the box, X can take the values 0, 1 and 2 and Y can take the values 0, 1, 2 and 3.

$P(X=0, Y=0)$ = P(drawing 3 balls none of which is white or red)

= P(all the 3 balls drawn are black)

$$= 4C_3/9C_3 = \frac{1}{21}$$

$P(X=0, Y=1)$ = P(drawing 1 red and 2 black balls)

$$= \frac{3C_1 \times 4C_2}{9C_3} = \frac{3}{14}$$

Similarly, $P(X=0, Y=2) = \dfrac{3C_2 \times 4C_1}{9C_3} = \dfrac{1}{7}$; $P(X=0, Y=3) = \dfrac{1}{84}$

$P(X=1, Y=0) = \dfrac{1}{7}$; $P(X=1, Y=1) = \dfrac{2}{7}$; $P(X=1, Y=2) = \dfrac{1}{14}$;

$P(X=1, Y=3) = 0$ (since only 3 balls are drawn)

$P(X=2, Y=0) = \dfrac{1}{21}$; $P(X=2, Y=1) = \dfrac{1}{28}$; $P(X=2, Y=2) = 0$;

$P(X=2, Y=3) = 0$

The joint probability distribution of (X, Y) may be represented in the form of a table as given below:

X	Y			
	0	1	2	3
0	$\frac{1}{21}$	$\frac{3}{14}$	$\frac{1}{7}$	$\frac{1}{84}$
1	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{14}$	0
2	$\frac{1}{21}$	$\frac{1}{28}$	0	0

Example:

For the bivariate probability distribution of (X, Y) given below, find $P(X \leq 1)$, $P(Y \leq 3)$, $P(X \leq 1, Y \leq 3)$, $P(X \leq 1/Y \leq 3)$, $P(Y \leq 3/X \leq 1)$ and $P(X + Y \leq 4)$.

X \ Y	1	2	3	4	5	6
0	0	0	1/32	2/32	2/32	3/32
1	1/16	1/16	1/8	1/8	1/8	1/8
2	1/32	1/32	1/64	1/64	0	2/64

Solution:

$$P(X \le 1) = P(X=0) + P(X=1)$$

$$= \sum_{j=1}^{6} P(X=0, Y=j) + \sum_{j=1}^{6} P(X=1, Y=j)$$

$$= \left(0+0+\frac{1}{32}+\frac{2}{32}+\frac{2}{32}+\frac{3}{32}\right)+\left(\frac{1}{16}+\frac{1}{16}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\right.$$

$$= \frac{1}{4}+\frac{5}{8}=\frac{7}{8}$$

$$P(Y \le 3) = P(Y=1) + P(Y=2) + P(Y=3)$$

$$= \sum_{i=0}^{2} P(X=i, Y=1) + \sum_{i=0}^{2} P(X=i, Y=2)$$

$$+ \sum_{i=0}^{2} P(X=i, Y=3)$$

$$= \left(0+\frac{1}{16}+\frac{1}{32}\right)+\left(0+\frac{1}{16}+\frac{1}{32}\right)+\left(\frac{1}{32}+\frac{1}{8}+\frac{1}{64}\right)$$

$$= \frac{3}{32}+\frac{3}{32}+\frac{11}{64}=\frac{23}{64}$$

P $\sum^{3}$ $\sum^{3}$

$$= \left(0+0+\frac{1}{32}\right)+\left(\frac{1}{16}+\frac{1}{16}+\frac{1}{8}\right)=\frac{9}{32}$$

$$P(X \le 1 / Y \le 3) = \frac{P(X \le 1, Y \le 3)}{P(Y \le 3)} = \frac{9/32}{23/64} = \frac{18}{23}$$

$$P(Y \le 3 / X \le 1) = \frac{P(X \le 1, Y \le 3)}{P(X \le 1)} = \frac{9/32}{7/8} = \frac{9}{28}$$

$$P(X+Y \le 4) = \sum_{j=1}^{4} P(X=0, Y=j) + \sum_{j=1}^{3} P(X=1, Y=j) + \sum_{j=1}^{2} P(X=2, Y=j)$$

$$= \frac{3}{32}+\frac{1}{4}+\frac{1}{16}=\frac{13}{32}$$

Exercise:

The joint probability mass function of (X, Y) is given by $p(x, y) = k(2x + 3y)$, $x = 0, 1, 2$; $y = 1, 2, 3$. Find all the marginal and conditional probability distributions. Also find the probability distribution of $(X + Y)$.

Example:

The joint pdf of a two-dimensional RV (X, Y) is given by $f(x, y) = xy^2 + \frac{x^2}{8}$, $0 \leq x \leq 2, 0 \leq y \leq 1$.

Compute $P(X > 1)$, $P\left(Y < \frac{1}{2}\right)$, $P(X > 1/Y < 1/2)$

$P\left(Y < \frac{1}{2}/X > 1\right)$, $P(X < Y)$ and $P(X + Y \leq 1)$.

Solution:

Here the rectangle defined by $0 \le x \le 2, 0 \le y \le 1$ is the range space R. $R_1, R_2, \ldots$,are event spaces.

(i) $P(X > 1) = \int\limits_{\substack{R_1 \\ (x>1)}} \int f(x, y)\, dx\, dy$

$$= \int_0^1 \int_1^2 \left(xy^2 + \frac{x^2}{8} \right) dxdy = \frac{19}{24}$$

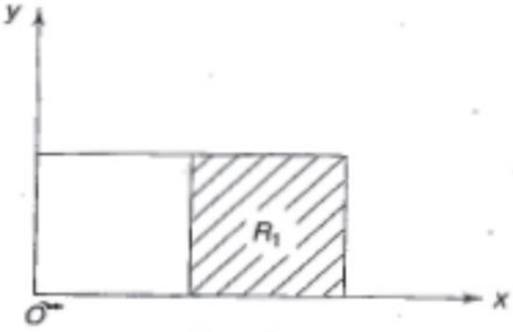

Fig. 2.1

(ii) $P(Y < 1/2) = \int\limits_{\substack{R_2 \\ \left(y<\frac{1}{2}\right)}} \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$

$$= \int_0^{1/2} \int_0^2 \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

$$= \frac{1}{4}$$

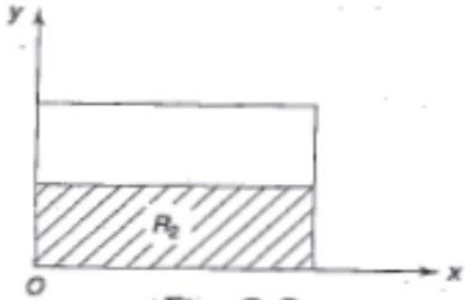

Fig. 2.2

(iii) $P(X > 1, Y < 1/2) = \underset{\substack{R_5 \\ \left(x>1 \,\&\, y<\frac{1}{2}\right)}}{\int} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^{1/2}\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$$

$$= \frac{5}{24}$$

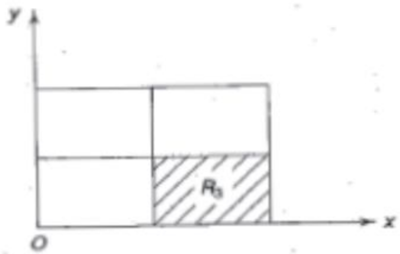

Fig. 2.3

(iv) $P(X > 1/Y < \frac{1}{2}) = \dfrac{P\left(X > 1, Y < \frac{1}{2}\right)}{P\left(Y < \frac{1}{2}\right)} = \dfrac{5/24}{1/4} = \dfrac{5}{6}$

(v) $P(Y < \frac{1}{2}/X > 1) = \dfrac{P\left(X > 1, Y < \frac{1}{2}\right)}{P(X > 1)} = \dfrac{5/24}{19/24} = \dfrac{5}{19}$

(vi) $P(X < Y) = \underset{\substack{R_4 \\ (x<y)}}{\int\int} \left(xy^2 + \frac{x^2}{8}\right) dx\, dy$

$$= \int_0^1\int_0^y \left(xy^2 + \frac{x^2}{8}\right) dx\, dy = \frac{53}{480}$$

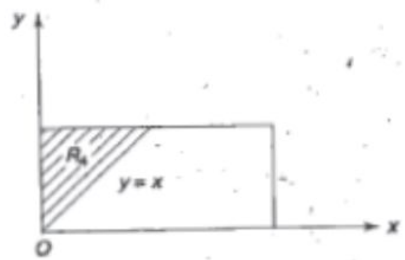

Fig. 2.4

(vii) $P(X+Y \leq 1) = \underset{\substack{R_5 \\ (x+y\leq 1)}}{\int\int} \left(xy^2 + \frac{x^2}{8}\right) \mathrm{d}x\, \mathrm{d}y$

$$= \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8}\right) \mathrm{d}x\, \mathrm{d}y = \frac{13}{480}$$

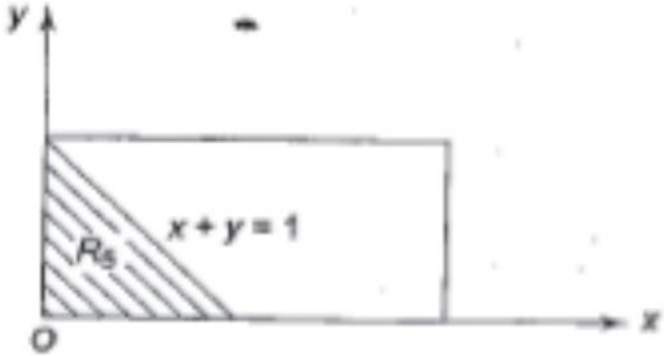

Fig. 2.5

Exercise:

Given $f_{XY}(x, y) = cx\,(x - y)$, $0 < x < 2$, $-x <, y < x$, and 0 elsewhere, (a) evaluate c, (b) find $f_X(x)$, (c) $f_{Y/X}\,(y/x)$ and (d) $f_Y(y)$.

Covariance

As the variance $E\{X - E(X)\}^2$ measures the variations of the R.V. X from its mean value $E(X)$, the quantity $E\{[X - E(X)]\,[Y - E(Y)]\}$ measures the simultaneous variation of two R.V.'s X and Y from their respective means and hence it is called *the covariance of* X, Y and denoted as Cov (X, Y).

Cov $(X, Y) = E\{[X - E(X)]\,[Y - E(Y)]\}$ is also called the *product moment* of X and Y and is also denoted as $p(X, Y)$.

$$\text{Cov}\,(X, Y) = \underline{\frac{1}{n}\Sigma x_i\, y_i - \frac{1}{n}\Sigma x_i \cdot \frac{1}{n}\Sigma y_i}$$

Moment Generating Function (MGF)

Moment Generating Function (MGF) of a RV X (discrete or continuous) is defind as $E(e^{tX})$, where t is a real variable and denoted as $M(t)$.

If X is discrete, then $M(t) = \sum_r e^{tx_r} p_r$,

where X takes the values $x_1, x_2, x_3, \ldots$, with probabilities $p_1, p_2, p_3, \ldots$

If X is a continuous RV with density function $f(x)$, then

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)\, dx$$

Properties of MGF

(Proofs of the properties are omitted, as the proofs of the corresponding properties of characteristic function will be given later.)

1. $M(t) = \sum_{n=0}^{\infty} t^n E(X^n)/\underline{|n}$

   i.e., $E(X^n) = \mu'_n$ is the co-efficient of $\frac{t^n}{\underline{|n}}$ in the expansion of $M(t)$ in series of powers of t.

2. $\mu'_n = E(X^n) = \left[\frac{d^n}{dt^n} M(t)\right]_{t=0}$.
3. If the MGF of X is $M_X(t)$ and if $Y = aX + b$, then $M_Y(t) = e^{bt} M_X(at)$.
4. If X and Y are independent RVs and $Z = X + Y$, then $M_Z(t) = M_X(t) M_Y(t)$.

Characteristic Function (CF)

Characteristic function of a RV X (discrete or continuous) is defined as $E(e^{i\omega X})$ and denoted as $\phi(\omega)$.

If X is a discrete RV that can take the values $x_1, x_2, \ldots,$ such that $P(X = x_r) = p_r$, then

$$\phi(\omega) = \sum_r e^{i\omega x_r} p_r$$

If X is a continuous RV with density function $f(x)$, then

$$\phi(\omega) = \int_{-\infty}^{\infty} e^{i\omega x} f(x)\, dx$$

Properties of Characteristic Function

1. $\mu'_n = E(X^n) =$ the coefficient of $\frac{i^n \omega^n}{n!}$ in the expansion of $\phi(\omega)$ in series of ascending powers of $i\omega$.

2. $\mu'_n = \frac{1}{i^n}\left[\frac{d^n}{d\omega^n}\phi(\omega)\right]_{\omega=0}$

3. If the characteristic function of a RV X is $f_X(\omega)$ and if $Y = aX + b$, then

$$\phi_Y(\omega) = e^{ib\omega}\,\phi_X(a\omega)$$

4. If X and Y are independent RVs, then

$$\phi_{X+Y}(\omega) = \phi_X(\omega) \times \phi_Y(\omega)$$