B.Sc. [ECS] - II

Sem - III

Probability Theory - II

PAGE NO.:

DATE: / /

Unit - I

Continuous Random Variable

Introduction:

We have seen that, there are three types of sample spaces i) finite ii) countably infinite and iii) continuous. In this chapter we discuss the random variables defined on continuous sample space.

A sample space which is finite or countably infinite is called as countable. If the sample space is not countable then it is called continuous. In other words for a continuous sample space Ω we can not have one-to-one correspondence between Ω and set of natural numbers $\{1, 2, ...\}$

eg. 1) Suppose weight of an oil bag having the capacity of 1 kg filled by an automatic filling machine is noted.

The sample space will be an interval in the neighbourhood of 1 kg such as $\Omega = (0.980, 1.005)$

2) Suppose is an experiment, life of an electronic component in hours is recorded.

The sample space in this case may be an interval as a part of R^+ such as $\Omega = (0, 5000)$

Note: A continuous sample space is a subset of Real line.

* Continuous Random Variable:

Def^n:- Define a random variable $X(\omega)$ as a real valued function on domain Ω. If the range set of $X(\omega)$ is continuous the r.v. is continous. The range set will be a subset of real line.

Examples of Continuous Random Variable

1) Weight of a person in kg.
2) Consumption of electricity of a town in a specific month.
3) Daily rainfall in cm. at a particular place.
4) life in hours of an electric component.

* Difference between continuous random variable and discrete random variable.

Continuous r.v.	Discrete r.v.
i> A continuous r.v. takes all posible values in a range set. The set is in the form of interval.	i> A discrete random variables takes only specific values.
ii> A continuous r.v. takes uncountably infinite values no probability mass can be attached to a particular value of r.v. X. Therefore $P(X=x)=0$ for all x.	ii> In case of discrete random variable probability mass is attached to individual values taken by random variable.

Continuous Probability Distribution :-

In case of discrete r.v. using the p.m.f. we get probability distribution of r.v. however in case of continuous r.v. probability mass is not attached to any particular value. It is attached to an interval. The probability attached to an interval depends upon its location.

e.g. $P(a<x<b)$ varies for different values of a and b. In order to obtain the probability associated with any interval, we need to take into account the concept of probability density.

A function $f(x)$ which is to be treated as a probability density function, should be a non-negative and continuous function of x. The probability that a variable x takes values in a small interval $\left(x-\frac{\delta x}{2}, x+\frac{\delta x}{2}\right)$ will be the product of length of interval and the value of density function $f(x)$ at the centre of interval.

$$\therefore p\left(x-\frac{\delta x}{2} < x < x+\frac{\delta x}{2}\right) \approx f(x)\cdot \delta x$$

Note: (1) Here we assume that the probability density is constant over the interval $\left(x-\frac{\delta x}{2}, x+\frac{\delta x}{2}\right)$.

This assumptions will not be valid for large interval. To overcome this difficulty we integrate $f(x)$ w.r.t. x over the given interval.

Thus, $P(a<x<b) = \int_a^b f(x)\,dx$

2) The above probability is a definite integral hence geometrically it is the area under curve $y=f(x)$ bounded by x axis and the ordinates at a and b.

Defn:- **Probability Density Function (p.d.f.):**

A real valued function $f(x)$ is called as probability density function (p.d.f.) of a continuous random variable x if

i) $f(x) \geq 0 \quad ; -\infty < x < \infty$

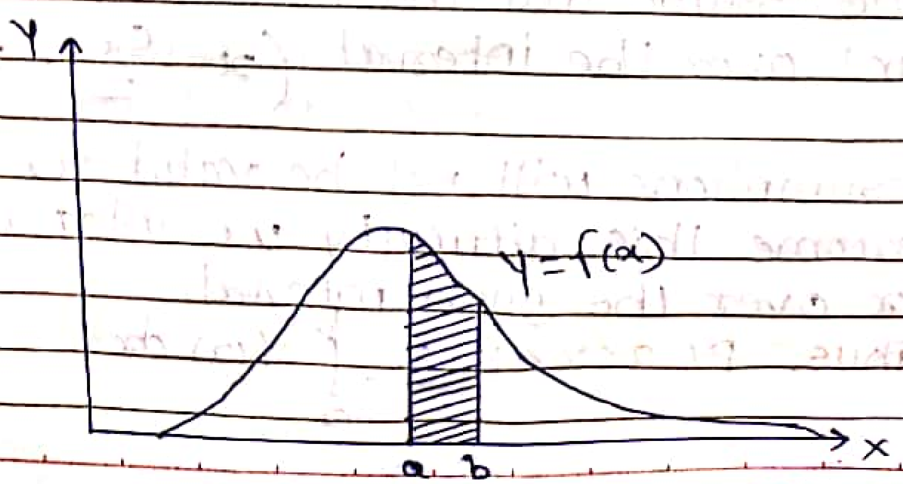
ii) $\int_{-\infty}^{\infty} f(x)\, dx = 1.$

Note: i) Since probability is associated with any individual value is zero.

$P(a<x<b) = P(a \leq x < b) = P(a < x \leq b) = P(a \leq x \leq b)$
$= \int_a^b f(x)\, dx =$ Area under the curve $f(x)$

bounded between x axis and the ordinates at a and b.

It is shown in following fig.



ii) $\int_{-\infty}^{\infty} f(x)\,dx$, can be interpreted as $P(-\infty < X < \infty)$

It also represents the total area under the density curve bounded by x-axis
Thus the total area under the curve is

$$P(-\infty < X < \infty) = \int_{-\infty}^{\infty} f(x)\,dx = 1$$

iii) If $A \subset R$, then $P(X \in A) = \int_A f(x)\,dx$

Ex: ① Verify the following functions are p.d.f.s or not.

i) $f(x) = 3x^2 \quad ; \quad 0 \le x \le 1$
$\quad = 0 \quad ; \quad$ otherwise

ii) $f(x) = 2e^{-x} \quad ; \quad x \ge 0$
$\quad = 0 \quad ; \quad$ otherwise

Sol:- To verify whether $f(x)$ is pdf we need to verify the following two conditions

ⓐ $f(x) \ge 0$, $\forall x$ and ⓑ $\int_{-\infty}^{\infty} f(x)\,dx = 1$

i) $f(x) = 3x^2 \ge 0 \quad \forall x$

and $\int_{-\infty}^{\infty} 3x^2\,dx = \int_0^1 3x^2\,dx = 3\int_0^1 x^2\,dx = \cancel{3}\frac{[x^3]_0^1}{\cancel{3}}$

$= 1$

ii) $f(x) = 2e^{-x} \ge 0 \quad \forall x$ and

$$\int_{-\infty}^{\infty} 2e^{-x}\,dx = 2\int_0^{\infty} e^{-x}\,dx = 2\left[\frac{e^{-x}}{-1}\right]_0^{\infty} = 2[0+1]$$

$= 2 \ne 1$

$\therefore$ Hence $f(x)$ is not a p.d.f.

* Distribution Function OR Cumulative Distribution Function (c.d.f.) :-

A distribution function is defined in case of a continuous r.v. analogus to that of discrete r.v. The summation (i.e. Σ) is to be replaced by integration (i.e. $\int$)

Def^n:- Let X be a continuous r.v. with p.d.f. $f(x)$. The distribution function or cumulative distribution function denoted by $F(x)$ is defined as.

$$F(x) = P(X \leq x) \quad ; -\infty < x < \infty$$
$$= \int_{-\infty}^{x} f(t)\,dt$$

Note: A r.v. X is defined to be continuous if $F(x)$ is continuous.

Ex: ① A r.v. X has p.d.f.

$$f(x) = 2e^{-2x} \quad ; \quad x > 0$$
$$= 0 \quad ; \quad \text{otherwise}$$

Find its distribution function

So^n:- By definition,

$$F(x) = \int_{-\infty}^{x} f(t)\,dt = \int_{0}^{x} 2e^{-2t}\,dt = 2\int_{0}^{x} e^{-2t}\,dt$$
$$= 2\left[\frac{e^{-2t}}{-2}\right]_0^x$$
$$= 2\left[\frac{e^{-2x}}{-2} - \frac{e^{-2(0)}}{-2}\right]$$

$$= 2\left[\frac{e^{-2x}}{-2} + \frac{e^{0}}{2}\right]$$

$$= 2\left[\frac{-e^{-2x}}{2} + \frac{1}{2}\right]$$

$$= \cancel{2}\left[\frac{-e^{-2x}+1}{\cancel{2}}\right]$$

$$= -e^{-2x}+1$$

$$\therefore F(x) = 1 - e^{-2x}$$

Since $F(x)$ is defined for all $x \in R$ we have to write,

$$F(x) = 0 \quad ; x < 0$$
$$= 1 - e^{-2x} \quad ; x \geq 0$$

Note: 1) We have seen above how to find d.f. given the p.d.f. We can also find p.d.f. from d.f. as follows

$$f(x) = \frac{d}{dx} F(x)$$

2) Given the d.f. $F(x)$ we can find $P(a < X \leq b)$ as follows

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

* Properties of Distribution Function:

1) Non-negative: $F(x)$ is a non-negative function. That is, $F(x) \geq 0 \ \forall x$.

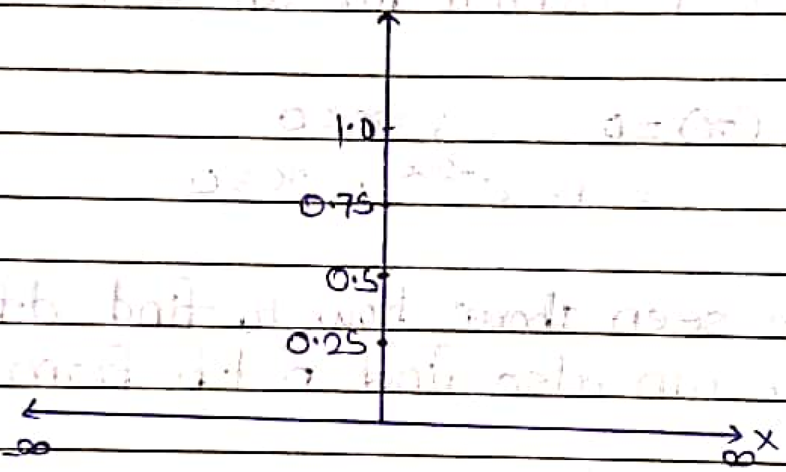
2) Non-decreasing: $F(x)$ is non-decreasing. That is if $a < b$ then $F(a) \leq F(b)$

3) $F(-\infty) = \lim_{x \to -\infty} F(x) = 0$ and

$F(\infty) = \lim_{x \to \infty} F(x) = 1$

4) Continuity: $F(x)$ is continuous for a continuous r.v. X

Graph of $F(x)$ of a continuous r.v. is a smooth continuous curve as shown in fig.



The above stated properties of distribution function are called as characteristic properties. It means every distribution function must satisfy these properties and any function satisfying these properties is a distribution function of some r.v.

EX: ① The life of tubelight (x in hours) follows the following distribution given by the p.d.f. $f(x)$

$$f(x) = \frac{K}{x^2} \quad ; \quad 1000 < x < 2000$$

$$= 0 \quad ; \quad \text{otherwise}$$

Determine the constant K and the distribution

$\int x^n = \frac{x^{n+1}}{n+1}$

PAGE NO.:
DATE: / /

function of x. Also compute the probability that a tubelight selected at random will have life between 1500 hours and 1600 hours.

Solⁿ:- Since $\int_{-\infty}^{\infty} f(x)\,dx = 1$

we have,

$$\int_{-\infty}^{\infty} \frac{k}{x^2}\,dx = 1$$

$$\int_{1000}^{2000} \frac{k}{x^2}\,dx = 1$$

$$\therefore k \int_{1000}^{2000} x^{-2}\,dx = 1$$

$$\therefore k \left[\frac{x^{-2+1}}{-2+1}\right]_{1000}^{2000} = 1$$

$$\therefore k \left[\frac{x^{-1}}{-1}\right]_{1000}^{2000} = 1$$

$$\therefore k \left[-\frac{1}{x}\right]_{1000}^{200} = 1$$

$$\therefore k \left[\frac{-1}{2000} + \frac{1}{1000}\right] = 1$$

$$\therefore k \left[\frac{-1}{2000} + \frac{1\times 2}{1000\times 2}\right] = 1$$

$$\therefore k \left[\frac{-1}{2000} + \frac{2}{2000}\right] = 1$$

$$\therefore k \left[\frac{-1+2}{2000}\right] = 1$$

$$\therefore k \left[\frac{1}{2000}\right] = 1$$

$$\therefore \frac{k}{2000} = 1$$

$\int t^n = \frac{t^{n+1}}{n+1}$

PAGE NO.:
DATE: / /

$\therefore k = 2000$

Distribution function $F(x)$ is given by.

$$F(x) = P[X \le x] = \int_{-\infty}^{\infty} f(t)\,dt$$

$$= \int_{1000}^{x} \frac{2000}{t^2}\,dt$$

$$= 2000 \int_{1000}^{x} \frac{1}{t^2}\,dt$$

$$= 2000 \int_{1000}^{x} t^{-2}\,dt$$

$$= 2000 \left[\frac{t^{-2+1}}{-2+1} \right]_{1000}^{x}$$

$$= 2000 \left[\frac{t^{-1}}{-1} \right]_{1000}^{x}$$

$$= 2000 \left[-\frac{1}{t} \right]_{1000}^{x}$$

$$= 2000 \left[-\frac{1}{x} + \frac{1}{1000} \right]$$

$$= \left[\frac{-2000}{x} + \frac{\cancel{2000}^{2}}{\cancel{1000}} \right]$$

$$= \left[\frac{-2000}{x} + 2 \right]$$

$$\therefore F(x) = 2 - \frac{2000}{x} \;;\; 1000 < x < 2000$$

Probability that the life of a tube is between 1500 and 1600 is given by.

$$F(1500 < X < 1600) = F(1600) - F(1500)$$

$$= \left(2 - \frac{2000}{1600}\right) - \left(2 - \frac{2000}{1500}\right)$$

$$= \left(\frac{1200}{1600}\right) - \left(\frac{1000}{1500}\right)$$

$$= \frac{3}{4} - \frac{2}{3}$$

$$= \frac{3\times3 - 4\times2}{4\times3}$$

$$= \frac{9-8}{12}$$

$$= \frac{1}{12}$$

$$\therefore F(1500 < x < 1600) = 0.0832$$

* Mean and variance:

We havens seen how to find expectation of discrete r.v. On similar lines it is defined for continuous r.v. The summation sign used in case of discrete r.v. is replaced by integration in case of continuous r.v.

Defn:- Let x be a continuous r.v. with p.d.f. f(x) then the mean or expectation of x is denoted by E(X) and is defined as,

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

Note: If the p.d.f. of x is defined (non-zero) over an interval [a,b] then

$$E(X) = \int_a^b x f(x)\, dx$$

* Expectation of a function of X :

If g(x) is a real valued function of a r.v. X then E[g(x)] is given by

$$E[g(x)] = \int_{-\infty}^{\infty} g(x) \cdot f(x)\, dx$$

Here also, whenever f(x) is defined over a finite interval [a,b] then

$$E[g(x)] = \int_a^b g(x) \cdot f(x)\, dx$$

* Variance:

Now we can define variance of x as follows

$$Var(X) = E[X - E(X)]^2$$

$$\therefore Var(X) = E(X^2) - [E(X)]^2$$

Ex: ① If X is a r.v. with p.d.f.

$$f(x) = 6x(1-x) \;;\; 0 < x < 1$$
$$= 0 \;;\; \text{otherwise}$$

find i) Mean ii) Variance of x.

Son:- Mean,

$$E(X) = \int_0^1 x f(x)\, dx$$

$$= \int_0^1 x \cdot 6x(1-x)\, dx$$

$$= \int_0^1 6x^2 - 6x^3 \, dx$$

$$= 6\int_0^1 x^2 \, dx - 6\int_0^1 x^3 \, dx$$

$$= {}^{2}\cancel{6}\left[\frac{x^3}{\cancel{3}_1}\right]_0^1 - {}^{3}\cancel{6}\left[\frac{x^4}{\cancel{4}_2}\right]_0^1$$

$$= 2\left[x^3\right]_0^1 - \frac{3}{2}\left[x^4\right]_0^1$$

$$= 2\left(1^3 - 0^3\right) - \frac{3}{2}\left(1^4 - 0^4\right)$$

$$= 2 - \frac{3}{2}$$

$$= \frac{4-3}{2}$$

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

$\therefore E(X) = 0.5$

$$Var(X) = E(X^2) - [E(X)]^2$$

where, $E(X^2) = \int_0^1 x^2 f(x) \, dx$

$$= \int_0^1 x^2 . 6x(1-x) \, dx$$

$$= \int_0^1 6x^3 - 6x^4 \, dx$$

$$= 6\int_0^1 x^3 \, dx - 6\int_0^1 x^4 \, dx$$

$$= 6\left[\frac{x^4}{4}\right]_0^1 - 6\left[\frac{x^5}{5}\right]_0^1$$

$$= \frac{6}{4}\left[x^4\right]_0^1 - \frac{6}{5}\left[x^5\right]_0^1$$

$$= \frac{3}{2}(1^4 - 0^4) - \frac{6}{5}(1^5 - 0^5)$$

$$= \frac{3}{2} - \frac{6}{5}$$

$$= \frac{5\times3 - 6\times2}{2\times5}$$

$$= \frac{15-12}{10}$$

$$= \frac{3}{10}$$

$\therefore E(X^2) = 0.3$

$$V(X) = E(X^2) - [E(X)]^2$$

$$= 0.3 - (0.5)^2$$

$$= 0.3 - 0.25$$

$\therefore V(X) = 0.05$

Scanned with CamScanner