UNIT – 1

CHAPTER- 1

RANDOM VARIABLES AND DISTRIBUTIONS

Introduction:

Observations are taken about the characteristic under study, which can take different values and hence referred to as a variable. The sample space of all the values of a variable can be (i) finite (ii) countably infinite or (iii) uncountably infinite. The variables with first two types of sample spaces are discrete random variables and the third type of sample space leads to continuous variable.

The value of the characteristic under study is known as a variable. Eg: age, demand, arrival time, service time, marks scored etc.

1. **Random Experiment :** An experiment having more than one outcome and the result of an experiment cannot be predicted in advance is called random experiment.
2. **Random Variable:** A real valued function associated with the outcome of the random experiment is called random variable (r.v.).

   . A **discrete random variable** is one whose set of assumed values is **countable** (arises from counting).

. A **continuous random variable** is one whose set of assumed values is uncountable (arises from measurement).

Examples :

(1) If the random experiment is throwing an unbiased dice, the random variable associated with this experiment can be defined as the number on the uppermost face, then the possible values of X are 1, 2, 3, 4, 5 and 6.

(2) The number of missed calls received on a particular day.

(3) Suppose oil contents in one litre bags are measured , then X = oil content in a bag, then X can take values 990 ml $\leq X \leq$ 1100 ml.

(4) Waiting time in a queue.

Notation: The random variable is denoted by uppercase letter and its value in lowercase letter. For e.g: X denotes a random variable while x denotes its value.

DISCRETE RANDOM VARIABLE:

If the number of possible values of that variable is finite or countably infinite, then the variable is called a discrete variable.

For e.g.:

(i) Age in completed years,

(ii) Number of arrivals at a clinic,

(iii) Number of jobs completed per day

(iv) Number of heads in 4 flips of a coin (possible outcomes are 0, 1, 2, 3, 4).

PROBABILITY MASS FUNCTION:

Let X is a discrete random variable. The possible values of X are denoted by the sample space S_x. Then the probability function

$P_x(x) = P(X = x) \quad x \in S_x$

Is known as **probability mass function (p.m.f.)**. This function satisfy the following two conditions:

(1) $P(x) \geq 0$ for all $x \in S_x$

(2) $\Sigma P(x) = 1$, summation over $x \in S_x$

CUMULATIVE PROBABILITY DISTRIBUTION FUNCTION (c.d.f.):

If X is a discrete random variable with probability mass function (p. m. f.) P (x) defined on sample space S, then the Cumulative (Probability) Distribution Function (c.d.f.) denoted by F(x) is defined as

$$F(x) = P(X \leq x) = \Sigma P(x_i) \quad x \in R \text{ and } i = -\alpha \text{ to } +\alpha$$

Properties of Cumulative probability distribution function:

(1) $0 \leq F(x) \leq 1$ for $x \in R$

(2) $F(-\alpha) = 0$ and $F(\alpha) = 1$

(3) F(x) is a non-decreasing function of x.

(4) F(x) is a step function and is continuous at right.

(5) If x_n and x_{n+1} are two consecutive values of the random variable X such that $x_n < x_{n+1,}$

then $P(x_{n+1}) = F(x_{n+1}) - F(x_n)$

(6) If a and b are two real numbers (a < b), then

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

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

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

(iv) $P(a < X < b) = F(b) - F(a) - P(X = b)$

(v) $P(X > a) = 1 - P(X \leq a) = 1 - F(a)$

(vi) $P(X < a) = F(a) - P(X = a)$

Median and Mode of the Distribution:

Median: Let X is a discrete random variable with Cumulative (Probability) Distribution Function (c.d.f.) F(x). Then the median M of the probability mass function of X is defined as the value of X such that $P(X \leq M) = P(X \geq M) = 0.5$

For calculation purpose we can define median as the first value of X for which $F(x) \geq 0.5$.

Mode: Let X be a discrete random variable with probability mass function (p.m.f.) $P_x(x)$ defined on sample space S. Then the mode of the probability mass function of X is defined as the value of X for which $P_x(x)$ is maximum.

CONTINUOUS RANDOM VARIABLE:

A Probability distribution is a function that describes the possible values of a random variable and their associated probabilities. A continuous random variable has a range in the form of an interval. Thus a continuous random variable assumes uncountably infinite values. Hence, if we try to attach probability mass to any point among infinitely many equally likely points, it becomes $1/\alpha=0$. Hence for a continuous random variable $P(X=x) = 0$. So in this case, probability is attached to an interval which is a subset of R, rather than a point. Since $P(X=x)$ is not defined, probability density i.e., probability of X taking value in a neighbourhood of x is considered. We consider a small interval of length dx around x such that dx->0 i.e.,$\{(x – dx/2) \leq X \leq (x + dx/2)\}$ and the probability of this interval is probability density attached to the value of X.

Probability Density Function(p.d.f.):

For a continuous random variable X with sample space S, if

$$P\{(x - dx/2) \leq X \leq (x + dx/2)\} = f(x)dx, \text{ for all } x \in S,$$

then f(x) is defined as Probability Density Function (p.f.d.) of random variable X and then $y = f(x)$ represents the equation of the probability curve for X.

If f(x) is the p.d.f. of x, then the probability that x belongs to A, where A is some interval, is given by the integral of f(x) over that interval, that is :

$$P(x \in A) = \int f(x)\, dx, \quad \text{integration over A.}$$

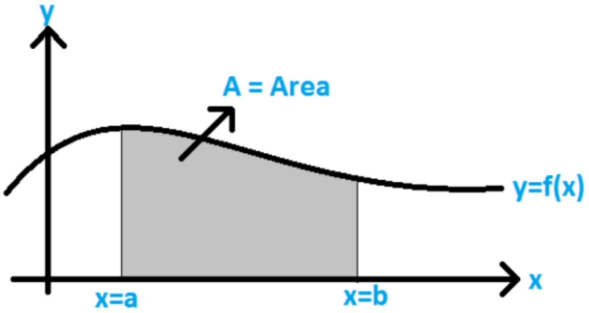
Properties of p.d.f.:

(1) $\int f(x)dx = 1$ (integration from -α to +α) implies that total area bounded by the curve of density function and X axis is equal to 1, when computed over entire range of variable X.

(2) $f(x) > 0$, for all $x \in S_x$

(3) Since the continuous random variable is defined over a continuous range of values , the graph of density function will also be continuous over that range.

(4) Since $y = f(x)$ is a continuous function of x, using calculus, area under $y = f(x)$ between $x = a$ and $x = b$ is $\int f(x)dx$, integration from a to b. Thus , for any (a,b) subset of S_x, we have
$P(a < X < b)$ = area under $y = f(x)$ between $x = a$ and $x = b = \int f(x)dx$.



(5) An implication of the fact that $P(X=x) = 0$ for all x when X is continuous is that one can be careless about the endpoints of intervals when finding probabilities of continuous random variables. That is:
$P(a < X < b) = P(a \le X < b) = P(a < X \le b) = P(a\le X\le b)$.
Note: $P(X=a) = P(a \le X \le a) = P(a < X < a)$ = area under $y = f(x)$ between $x = a$ and $x = a$. (using property 4.)

Cumulative (probability) distribution Function (c.d.f.):

If X is a continuous random variable with probability density function (p.d.f.) f(x) defined on sample space S, then Cumulative (Probability) Distribution Function (c.d.f.) denoted by F(x) is defined as

$F(x) = P(X \leq x) = \int f(t)dt = \int f(t)dt$ (integration from $-\alpha$ to x) = Area to the left of x bounded by the curve $y = f(x)$

Note: In discrete random variables, F(x) is a non-decreasing step function. For continuous random variables, F(x) is a non-decreasing continuous function.

Let sample space for X be $S_x = \{x/ a < x < b\}$.

Consider $F(x) = \int f(t)dt$ $(-\alpha$ to $x) = \int f(t)dt$ $(-\alpha$ to $a)$ $+ \int f(t)dt$ (a to x), but $f(t) = 0$ for $t \in (-\alpha, a)$. Hence c.d.f. can be calculated by integrating p.d.f. over the range (a,x) i.e., (lower limit of S_x).

The cumulative distribution function for continuous random variable is just a straightforward extension of that of the discrete case. Only summation sign is replaced by an integral.

Properties of c.d.f.:

(1) F(x) is a non-decreasing continuous function , i.e., if $a < b \Rightarrow F(a) < F(b)$

(2) Since F(x) is continuous function of x, it is differentiable, hence d/dx[F(x)] = f(x) for all x ϵ S_x.

(3) If a, b ϵ S_x, then P(a < x < b) = F(b) – F(a)

(4) Since F(x) = ∫ f(t)dt [-α to x]

(i) F(-α) = 0

Let lower limit of X = a, then F(a) = 0

(ii) F(α) = ∫ f(t) dt = 1 [-α to α]

Let b = upper limit of x, then F(b) = 1.

MATHEMATICAL EXPECTATION:

(a) If X is a discrete random variable with the sample space S_x and probability mass function P(x), then the expectation of g(X) is given by

E[g(X)] = Σ g(x) * P(x)

(b) Let X be a continuous random variable with probability density function (p.d.f.) f(x) defined on sample space S, then

(i) Expectation of a function : If g(x) is any function of x, then,

E[g(x)] = ∫ g(x)f(x) dx, [integration over S_x].

If g(x) = X, then we get E(X) which is known as mean of the random variable X.

Mean = E(X) = Σ X * P(X) if X is a discrete variable.

Mean = $E(X) = \int xf(x)\ dx$ if X is a continuous random variable.

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

Properties:

(1) $E(aX + b) = aE(X) + b$

Here $g(X) = aX + b$

By definition, $E[g(X)] = \Sigma g(x) * P(x)$

$$
\begin{aligned}
E(aX + b) &= \Sigma\ (ax + b) * P(x) \\
&= a\Sigma xP(x) + bP(x) \\
&= aE(X) + b \qquad [\text{since } \Sigma\ P(x) = 1]
\end{aligned}
$$

In case of continuous r.v., summation will be replaced by integration.

Particular cases:

(i) If $b = 0$, then $E(aX) = aE(X)$

(ii) If $a = 0$, then $E(b) = b$

(iii) If $a = 1$, then $E(X + b) = E(X) + b$

(2) $V(aX + b) = a^2V(X)$

$Y = aX + b$

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

$Y – E(Y) = a[X – E(X)]$

So, $E[Y – E(Y)]^2 = a^2[X – E(X)]^2$

$\Rightarrow V(Y) = a^2V(X)$

$\Rightarrow V(aX + b) = a^2V(X)$

MOMENTS:

(a) The rth raw moment about origin 0 is given by $\mu_r^{\prime} = E(X^r) = \Sigma\ x^rP(x)$ if X is a discrete random variable

$= \int x^rf(x)dx$, if X is a continuous random variable.

(b) The rth raw moment about ‘a’ $\mu_{r(a)} = E(X – a)^r$

$\mu_{r(a)}^{\prime} = E(X – a)^r = \Sigma\ (x – a)^rP(x)$, if X is a discrete random variable.

$= \int (x – a)^rf(x)dx$, if X is a continuous random variable.

(c) The rth central moment:

$\mu_r = E(X - \mu_r^{\prime})^r = \Sigma\ (X - \mu_r^{\prime})^rP(x)$, if X is a discrete random variable.

$= \int (X - \mu_r^{\prime})^rf(x)dx$, if X is a continuous random variable.

Relations between Central Moments and Raw Moments:

$\mu_r' = \Sigma\, x^r P(x) \quad r = 1, 2, 3, 4$

$\mu_r = \Sigma\, (X - \mu_r')^r P(x) \quad r = 1, 2, 3, 4$

When r = 0, $\boldsymbol{\mu_0' = 1, \quad \mu_0 = 1}$

When r = 1, $\boldsymbol{\mu_1' = \mu, \quad \mu_1 = 0}$

$\mu_1' = \Sigma\, xP(x) = E(X) = \mu$

$\mu_1 = \Sigma\, (x - \mu_1')P(x) = \Sigma xP(x) - \mu_1' \Sigma P(x)$

$= \mu_1' - \mu_1' = 0$

When r = 2, $\boldsymbol{\mu_2 = \mu_2' - (\mu_1')^2}$

$\mu_2 = \Sigma\, (x - \mu_1')^2 P(x)$

$= \Sigma\, x^2P(x) - 2\, \mu_1' \Sigma xP(x) + (\mu_1')^2\, \Sigma P(x)$

$= \mu_2' - (\mu_1')^2$

When r = 3, $\boldsymbol{\mu_3 = \mu_3' - 3\, \mu_2' \mu_1' + 2(\mu_1')^3}$

$\mu_3 = \Sigma\, (x - \mu_1')^3 P(x)$

$= \Sigma x^3 P(x) - 3\, \mu_1' \Sigma x^2 P(x) + 3(\mu_1')^2\, \Sigma\, xP(x) - (\mu_1')^3\, \Sigma P(x)$

$= \mu_3' - 3\, \mu_1'\, \mu_2' + 3(\mu_1')^2\, \mu_1' - (\mu_1')^3$

$= \mu_3' - 3\, \mu_1'\, \mu_2' + 2(\mu_1')^3$

When r = 4, $\boldsymbol{\mu_4 = \mu_4' - 4\, \mu_3' \mu_1' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4}$

Similarly, we can prove that the results for continuous random variable.

In general,

$$\mu_r = \mu_r{}' - ({}^rC_1)\mu'_{r-1}\mu'_1 + ({}^rC_2)\mu'_{r-2}(\mu_1')^2 - ({}^rC_3)\,\mu'_{r-3}(\mu_1')^3 + \ldots\ldots. \quad (-1)^{r+1}[({}^rC_{r-1}) - 1]\,(\mu_1')^r.$$

SKEWNESS AND KURTOSIS:

Skewness means lack of symmetry of tails (about mean) of a probability distribution curve.

Karl Pearson’s measure of skewness in terms of moments is given by

$$\mathbf{B_1 = \mu_3^2/\mu_2^3} \quad \text{and}$$

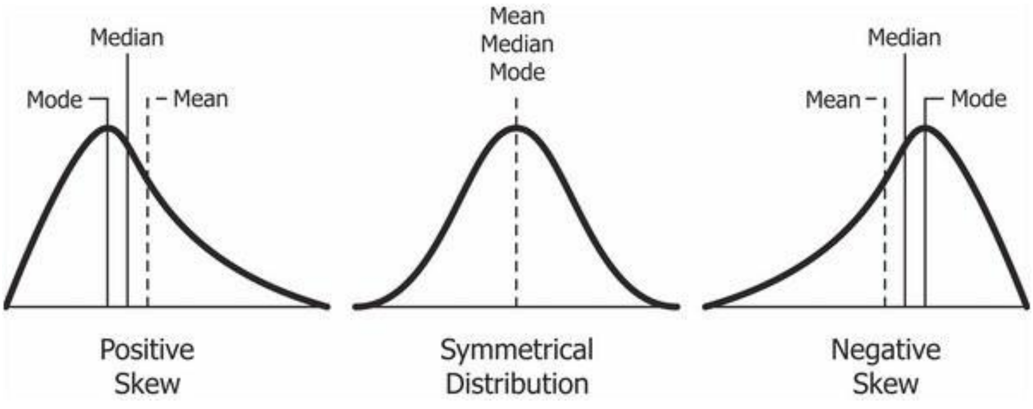
The direction of Skewness is measured by $\boldsymbol{\gamma_1 = \pm\sqrt{B_1}}$

Where sign of γ_1 is sign of μ_2.

(i) If $\gamma_1 > 0$, then the probability distribution is positively skewed curve.

(ii) If $\gamma_1 = 0$, then the probability distribution is a symmetric curve.

(iii) If $\gamma_1 < 0$, then the probability distribution is negatively skewed curve.



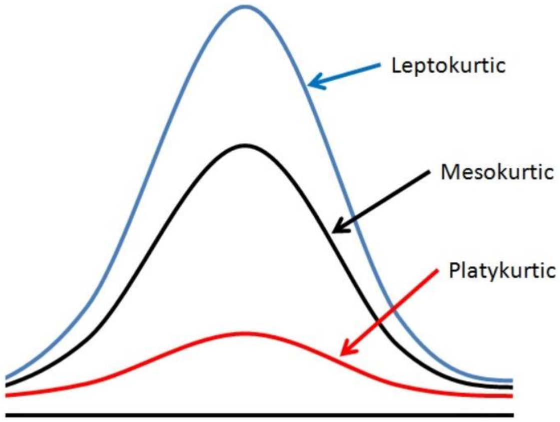
KURTOSIS: It relates to peakedness of probability distribution. It is measured in terms of

$$\mathbf{B_2 = \mu_4/\mu_2^2 \quad and\ \gamma_2 = B_2 - 3}$$

(i) If $\beta_2 < 3$ or $\gamma_2 < 0$ then probability distribution is **Platykurtic.**

(ii) If $\beta_2 > 3$ or $\gamma_2 > 0$ then probability distribution is **Leptokurtic.**

(iii) If $\beta_2 = 3$ or $\gamma_2 = 0$ then probability distribution is **Mesokurtic.**


Leptokurtic
Mesokurtic
Platykurtic