UNIT-III

Expectation:

Definition: - Let X be a r.v. with pdf(p.m.f) $f(x)$. Then its mathematical expectation is given by

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad \text{for continuous r.v.}$$

$$= \sum_x x f(x) dx \quad \text{for discrete r.v.}$$

provided the RHS integral (or) series is absolutely convergent.

i.e., $\int |x| f(x) dx < \infty$ (or) $\sum_x |x| f(x) < \infty$.

Expectation of a function of a R.V.

Let X be a r.v with pdf (pmf) $f(x)$ and D.F $F(x)$. If $g(\cdot)$ is a function & $g(x)$ is a r.v and $E[g(x)]$ exists, then the expectation of the function $g(x)$ is given by

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\,(x) = \int_{-\infty}^{\infty} g(x) f(x) dx \quad \text{for continuous case}$$

$$= \sum_x g(x) f(x) \quad \text{for discrete case}$$

By defn $E[g(x)] = E(y) = \int y \, d H_y(y)$ where $Y = g(x)$

$$= \int y \, h(y) dy \quad \text{(for continuous case)}$$

$$= \sum_y y \, h(y) \quad \text{(for discrete case)}.$$

Note:

If $g(x) = x^r$ (r: +ve integer),

then $E(x^r) = \int_{-\infty}^{\infty} x^r f(x) dx = \mu_r'$, the r^{th} moment about origin.

If $r=1$, $E(x) = \int x f(x) dx = \mu_1' =$ Mean

If $r=2$, $E(x^2) = \int x^2 f(x) dx = \mu_2'$

and $\mu_2 = \mu_2' - (\mu_1')^2 = E(x^2) - [E(x)]^2 =$ Variance of x.

1) Addition Theorem on Expectation:

For 2 variables: If X & Y be continuous r.v's with joint then

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

provided all the Expectations Exist.

For 'n' variables:

If $X_1, X_2, \dots X_n$ are random variables then

$$E(X_1+X_2+\dots+X_n) = E(X_1)+E(X_2)+\dots+E(X_n)$$

if all the expectations Exist.

2.) Multiplication Theorem on Expectation:

For 2 variables:-

If X & Y be two independent r.v's, then $E(XY) = E(X).E(Y)$.

For 'n' variables:

If $X_1, X_2, \dots X_n$ are 'n' independent r.v's, then

$$E(X_1.X_2.\dots X_n) = E(X_1).E(X_2)\dots E(X_n)$$

provided all the Expectations Exist.

3.) If X is a r.v and a is a constant, then

i) $E[a.\psi(x)] = a.E[\psi(x)]$

ii) $E[\psi(x)+a] = E[\psi(x)]+a$

where $\psi(x)$ is a function of X, and if all the Expectations Exist.

4) If X is a r.v and a & b are constants, then

$$E(aX+b) = a.E(X)+b$$

5) Let $X_1, X_2, \dots X_n$ be any 'n' r.v's and if $a_1, a_2, \dots a_n$ are constants then

$$E\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n} a_i E(X_i)$$

6) If $X \geq 0$ and $Y \geq 0$ are two r.v's then

i) $E(X \geq 0)$

ii) $E(Y) \leq E(X)$ if $Y \leq X$

iii) $|E(X)| = E|X|$.

7) If μ_r' exists, then μ_s' exists for all $1 \le s \le r$.

i.e., if $E(x^r) < \infty$, then $E(x^s)$ exists $\forall$ $1 \le s \le r$

8) If x is a r.v, then $V(ax+b) = a^2 V(x)$
where a & b are constants.

9) Covariance:

If X & Y are two r.v's and a & b are constants, then

i) Covariance between X & Y is defined as

$$Cov(X,Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$
$$= E(XY) - E(X).E(Y) \quad \text{if } X \text{ \& } Y \text{ are not independent}$$
$$= 0 \quad \text{if } X \text{ \& } Y \text{ are independent i.e. } E(XY) = E(X).E(Y).$$

ii) $Cov(X+a, Y+b) = Cov(X,Y)$

iii) $Cov(aX, bY) = ab\, Cov(X,Y)$

iv) $Cov\left(\frac{X-\bar{X}}{\sigma_X}, \frac{Y-\bar{Y}}{\sigma_Y}\right) = \frac{1}{\sigma_X \sigma_Y} Cov(X,Y).$

v) $Cov(aX+b, cY+d) = a.c\, Cov(X,Y)$ where c & d are constants

vi) $Cov(X+Y, Z) = Cov(X,Z) + Cov(Y,Z)$ where Z is also a r.v.

vii) $Cov(aX+bY, cX+dY)$
$$= ac\,\sigma_X^2 + bd\,\sigma_Y^2 + (ad+bc)\, Cov(X,Y).$$

10) Correlation Coefficient:

The Correlation Coefficient between X & Y is

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E[\{X-E(X)\}\{Y-E(Y)\}]}{E[X-E(X)]^2\, E[Y-E(Y)]^2}$$
$$= \frac{E(XY) - E(X).E(Y)}{\sigma_X \sigma_Y}.$$

11) If X & Y are independent r.v's then

$$E[h(X).k(Y)] = E[h(X)].E[k(Y)]$$

where $h(.)$ is a function of X alone & $k(.)$ is a function of Y alone provided all the expectations exist.

Source: Fundamentals of Mathematical Statistics by V. K. Kapoor & S. C. Gupta

Conditional Expectation

Discrete Case. The conditional expectation or mean value of a continuous function $g(X, Y)$ given that $Y = y_j$, is defined by

$$E\{g(X,Y) \mid Y = y_j\} = \sum_{i=1}^{\infty} g(x_i, y_j)\, P(X = x_i \mid Y = y_j)$$

$$= \frac{\sum_{i=1}^{\infty} g(x_i, y_j)\, P(X = x_i \cap Y = y_j)}{P(Y = y_j)} \qquad \text{...(6·46)}$$

i.e., $E[g(X,Y) \mid Y = y_j]$ is nothing but the expectation of the function $g(X, y_j)$ of X w.r.t. the conditional distribution of X when $Y = y_j$.

In particular, the conditional expectation of a discrete random variable X given $Y = y_j$ is

$$E(X \mid Y = y_j) = \sum_{i=1}^{\infty} x_i\, P(X = x_i \mid Y = y_j) \qquad \text{...(6·47)}$$

The conditional variance of X given $Y = y_j$ is likewise given by

$$V(X \mid Y = y_j) = E[\{X - E(X \mid Y = y_j)\}^2 \mid Y = y_j] \qquad \text{...(6·47a)}$$

The conditional expectation of $g(X, Y)$ and the conditional variance of Y given $X = x_i$ may also be defined in an exactly similar manner.

Continuous Case. The conditional expectation of $g(X, Y)$ on the hypothesis $Y = y$ is defined by

$$E\{g(X,Y) \mid Y = y\} = \int_{-\infty}^{\infty} g(x, y)\, f_{X|Y}(x \mid y)\, dx$$

$$= \frac{\int_{-\infty}^{\infty} g(x, y)\, f(x, y)\, dx}{f_Y(y)} \qquad \text{...(6·48)}$$

In particular, the conditional mean of X given $Y = y$ is defined by

$$E(X \mid Y = y) = \frac{\int_{-\infty}^{\infty} x\, f(x, y)\, dx}{f_Y(y)}$$

Similarly, we define

$$E(Y \mid X = x) = \frac{\int_{-\infty}^{\infty} y\, f(x, y)\, dy}{f_X(x)} \qquad \text{...(6·48 a)}$$

The conditional variance of X may be defined as

$$V(X \mid Y = y) = E\left[\{X - E(X \mid Y = y)\}^2 \mid Y = y\right]$$

Similarly, we define

$$V(Y \mid X = x) = E\left[\{Y - E(Y \mid X = x)\}^2 \mid X = x\right] \qquad \text{...(6·49)}$$

Theorem 6·13. *The expected value of X is equal to the expectation of the conditional expectation of X given Y. Symbolically,*

$$E(X) = E[E(X \mid Y)] \qquad \ldots(6{\cdot}50)$$

[Calicut Univ. B.Sc. (Main Stat.), 1980]

Proof. $E[E(X \mid Y)] = E[\sum_i x_i\ P(X = x_i \mid Y = y_j)]$

$$= E\left[\sum_i x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right]$$

$$= \sum_j \left[\sum_i \left\{x_i \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right\}\right] P(Y = y_j)$$

$$= \sum_j \sum_i x_i\ P(X = x_i \cap Y = y_j)$$

$$= \sum_i \left[x_i \{\sum_j P(X = x_i \cap Y = y_j)\}\right]$$

$$= \sum x_i P(X = x_i) = E(X).$$

Theorem 6·14. *The variance of X can be regarded as consisting of two parts, the expectation of the conditional variance and the variance of the conditional expectation.* Symbolically,

$$V(X) = E[V(X \mid Y)] + V[E(X \mid Y)] \qquad \ldots(6{\cdot}51)$$

Proof. $E[V(X \mid Y)] + V[E(X \mid Y)]$

$$= E\left[E(X^2 \mid Y) - \{E(X \mid Y)\}^2\right] + E[\{E(X \mid Y)\}^2] - [E\{E(X \mid Y)\}]^2$$

$$= E[E(X^2 \mid Y)] - E[\{E(X \mid Y)\}^2] + E[\{E(X \mid Y)\}^2] - [E\{E(X \mid Y)\}]^2$$

$$= E[E(X^2 \mid Y)] - [E(X)]^2 \qquad (c.f.\ \text{Theorem } 6{\cdot}13)$$

$$= E[\sum_i x_i^2 P(X = x_i \mid Y = y_j)] - [E(X)]^2$$

$$= E\left[\sum_i x_i^2 \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right] - [E(X)]^2$$

$$= \sum_j \left\{\left[\sum_i x_i^2 \frac{P(X = x_i \cap Y = y_j)}{P(Y = y_j)}\right] P(Y = y_j)\right\} - [E(X)]^2$$

$$= \sum_i \left[x_i^2 \sum_j P(X = x_i \cap Y = y_j)\right] - [E(X)]^2$$

$$= \sum_i x_i^2\ P(X = x_i) - [E(X)]^2$$

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

Hence the theorem.

Theorem 6·15. *Let A and B be two mutually exclusive events, then*

$$E(X \mid A \cup B) = \frac{P(A)\,E(X \mid A) + P(B)\,E(X \mid B)}{P(A \cup B)}$$

where by def.,

$$E(X \mid A) = \frac{1}{P(A)} \sum_{x_i \in A} x_i\, P(X = x_i)$$

Proof. $$E(X \mid A \cup B) = \frac{1}{P(A \cup B)} \sum_{x_i \in A \cup B} x_i\, P(X = x_i)$$

Since A and B are mutually exclusive events,

$$\sum_{x_i \in A \cup B} x_i\, P(X = x_i) = \sum_{x_i \in A} x_i\, P(X = x_i) + \sum_{x_i \in B} x_i\, P(X = x_i)$$

$$\therefore\; E(X \mid A \cup B) = \frac{1}{P(A \cup B)} \left[P(A)\,E(X \mid A) + P(B)\,E(X \mid B)\right]$$

Source: NOTES ON PROBABILITY by Greg Lawler. Last Updated: March 21, 2016

Properties of Conditional Expectation

- $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}]] = \mathbb{E}[X]$.

  Proof. This is a special case of (16) where $A = \Omega$.

- If $a, b \in \mathbb{R}$, $\mathbb{E}[aX + bY \mid \mathcal{G}] = a\,\mathbb{E}[X \mid \mathcal{G}] + b\,\mathbb{E}[Y \mid \mathcal{G}]$. (This equality and others below are really equalities up an event of probability zero.)

Proof. Note that $a\,\mathbb{E}[X \mid \mathcal{G}] + b\,\mathbb{E}[Y \mid \mathcal{G}]$ is $\mathcal{G}$-measurable. Also, if $A \in \mathcal{G}$,

$$\begin{aligned}
\int_A (a\,\mathbb{E}[X \mid \mathcal{G}] + b\,\mathbb{E}[Y \mid \mathcal{G}])\, d\mathbb{P} &= a\int_A \mathbb{E}[X \mid \mathcal{G}]\, d\mathbb{P} + b\int_A \mathbb{E}[Y \mid \mathcal{G}]\, dP\mathbb{P} \\
&= a\int_A X\, d\mathbb{P} + b\int_A Y\, d\mathbb{P} \\
&= \int_A (aX + bY)\, d\mathbb{P}.
\end{aligned}$$

The result follows by uniqueness of the conditional expectation.

- If X is $\mathcal{G}$-measurable, then $\mathbb{E}[X \mid \mathcal{G}] = X$.
- If $\mathcal{G}$ is independent of X, then $\mathbb{E}[X \mid \mathcal{G}] = \mathbb{E}[X]$. ($\mathcal{G}$ is independent of X if $\mathcal{G}$ is independent of the σ-algebra generated by X. Equivalently, they are independent if for every $A \in \mathcal{G}$, 1_A is independent of X.)

  Proof. The constant random variable $\mathbb{E}[X]$ is certainly $\mathcal{G}$-measurable. Also, if $A \in \mathcal{G}$,

$$\int_A \mathbb{E}[X]\, d\mathbb{P} = \mathbb{P}(A)\,\mathbb{E}[X] = \mathbb{E}[1_A]\,\mathbb{E}[X] = \mathbb{E}[1_A X] = \int_A X\, d\mathbb{P}.$$

- If $\mathcal{H} \subset \mathcal{G} \subset \mathcal{F}$ are σ-algebras, then

$$\mathbb{E}[X \mid \mathcal{H}] = \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}]. \tag{17}$$

Proof. Clearly, $\mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}]$ is $\mathcal{H}$-measurable. If $A \in \mathcal{H}$, then $A \in \mathcal{G}$, and hence

$$\int_A \mathbb{E}[\mathbb{E}[X \mid \mathcal{G}] \mid \mathcal{H}] \, d\mathbb{P} = \int_A \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_A X \, d\mathbb{P}.$$

- If Y is $\mathcal{G}$-measurable, then

$$\mathbb{E}[XY \mid \mathcal{G}] = Y \, \mathbb{E}[X \mid \mathcal{G}]. \tag{18}$$

Proof. Clearly $Y \, \mathbb{E}[X \mid \mathcal{G}]$ is $\mathcal{G}$-measurable. We need to show that for every $A \in \mathcal{G}$,

$$\int_A Y \, \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_A YX \, d\mathbb{P}. \tag{19}$$

We will first consider the case where $X, Y \geq 0$. Note that this implies that $\mathbb{E}[X \mid \mathcal{G}] \geq 0$ a.s., since $\int_A \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} = \int_A X \, d\mathbb{P} \geq 0$ for every $A \in \mathcal{G}$. Find simple $\mathcal{G}$-measurable random variables $0 \leq Y_1 \leq Y_2 \leq \cdots$ such that $Y_n \to Y$. Then $Y_n X$ converges monotonically to YX and $Y_n \, \mathbb{E}[X \mid \mathcal{G}]$ converges monotonically to $Y \, \mathbb{E}[X \mid \mathcal{G}]$. If $A \in \mathcal{G}$ and

$$Z = \sum_{j=1}^{n} c_j 1_{B_j}, \quad B_j \in \mathcal{G},$$

is a $\mathcal{G}$-measurable simple random variable,

$$\begin{aligned}
\int_A Z \, \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} &= \sum_{j=1}^{n} c_j \int_A 1_{B_j} \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} \\
&= \sum_{j=1}^{n} c_j \int_{A \cap B_j} \mathbb{E}[X \mid \mathcal{G}] \, d\mathbb{P} \\
&= \sum_{j=1}^{n} c_j \int_{A \cap B_j} X \, d\mathbb{P} \\
&= \int_A [\sum_{j=1}^{n} c_j 1_{B_j}] X \, d\mathbb{P} \\
&= \int_A ZX \, d\mathbb{P}.
\end{aligned}$$

Hence (19) holds for simple nonnegative Y and nonnegative X, and hence by the monotone convergence theorem it holds for all nonnegative Y and nonnegative X. For general Y, X, we write $Y = Y^+ - Y^-, X = X^+ - X^-$ and use linearity of expectation and conditional expectation.

If X, Y are random variables and X is integrable we write $\mathbb{E}[X \mid Y]$ for $\mathbb{E}[X \mid \sigma(Y)]$, Here $\sigma(Y)$ denotes the σ-algebra generated by Y. Intuitively, $\mathbb{E}[X \mid Y]$ is the best guess of X given the value of Y. Note that $\mathbb{E}[X \mid Y]$ can be written as $\phi(Y)$ for some function ϕ, i.e., for each possible value of Y there is a value of $\mathbb{E}[X \mid Y]$. Elementary texts often write this as $\mathbb{E}[X \mid Y = y]$ to indicate that for each value of y there is a value of $\mathbb{E}[X \mid Y]$. Similarly, we can define $\mathbb{E}[X \mid Y_1, \ldots Y_n]$.

Inequalities Based on Expectation

HOLDER'S INEQUALITY:

$$E|XY| \le [E|X|^r]^{1/r} \cdot E[|Y|^s]^{1/s} \text{ where } r>0 \ \& \ \frac{1}{r}+\frac{1}{s}=1.$$

Proof:

Consider $\varphi(p) = \frac{p^r}{r} + \frac{p^{-s}}{s}$; $0<p<\infty$

Suppose $\varphi(p)$ has minimum at $p=1$, then $\varphi(1) = \frac{1}{r}+\frac{1}{s}=1$

Let $p = \frac{a^{1/s}}{b^{1/r}}$, $a>0$, $b>0$.

Then $\varphi(p) = \left[\frac{a^{1/s}}{b^{1/r}}\right]^r \cdot \frac{1}{r} + \left[\frac{a^{1/s}}{b^{1/r}}\right]^{-s} \cdot \frac{1}{s}$

$$= \frac{a^{r/s}}{b \cdot r} + \frac{a^{-1}}{b^{-s/r}} \cdot \frac{1}{s}$$

$$= (br)^{-1} \cdot a^{r/s} + (as)^{-1} \cdot b^{s/r} \ge 1$$

Multiply by ab, on both sides

$$r^{-1} \cdot a^{(r/s)+1} + s^{-1} b^{(s/r)+1} \ge ab.$$

$$\Rightarrow \quad r^{-1} a^{\frac{r+s}{s}} + s^{-1} b^{\frac{s+r}{r}} \ge ab \quad \text{——— (1)}$$

As $1/r + 1/s = 1$, $\frac{r+s}{sr} = 1$

$$\Rightarrow \frac{r+s}{r} = s \ \& \ \frac{r+s}{s} = r. \quad \text{——— (2)}$$

Using (2), (1) can be written as

$$r^{-1} a^r + s^{-1} b^r \ge ab \quad \text{——— (3)}$$

Let $a = \frac{|X(\omega)|}{[E|X|^r]^{1/r}}$ and $b = \frac{|Y(\omega)|}{[E|Y|^s]^{1/s}}$

Now (3) becomes,

$$r^{-1}\left[\frac{|X|}{[E|X|^r]^{1/r}}\right]^r + s^{-1}\left[\frac{|Y|}{[E|Y|^s]^{1/s}}\right]^s \ge \frac{|X| \cdot |Y|}{[E|X|^r]^{1/r} \cdot [E|Y|^s]^{1/s}}$$

$$\Rightarrow \frac{1}{r} \cdot \frac{|X|^r}{E|X|^r} + \frac{1}{s} \cdot \frac{|Y|^s}{E|Y|^s} \ge \frac{|X| \cdot |Y|}{[E|X|^r]^{1/r} \cdot [E|Y|^s]^{1/s}}$$

$$\Rightarrow \quad \frac{s\,|X|^r E|Y|^s + r\,|Y|^s E|X|^r}{rs\, E|X|^r E|Y|^s} \geq \frac{|X|\cdot|Y|}{\left[E|X|^r\right]^{1/r} \cdot \left[E|Y|^s\right]^{1/s}}$$

Taking Expectation, we get

$$\frac{s+r}{rs} \geq \frac{E|X|\cdot E|Y|}{\left[E|X|^r\right]^{1/r} \cdot \left[|Y|^s\right]^{1/s}}$$

$$\Rightarrow E|XY| \leq \left[E|X|^r\right]^{1/r} \cdot \left[E|Y|^s\right]^{1/s} \quad \left[\because \frac{s+r}{rs} = 1\right]$$

BASIC INEQUALITY:

Let X be an arbitrary r.v and g on R be a non-negative Borel function. If g is even and non-decreasing on $[0, \infty]$ then for every $a > 0$

$$\frac{E[g(X)] - g(a)}{a.s\ \operatorname{Sup} g(x)} \leq P[|X| \geq a] \leq \frac{E[g(X)]}{g(a)} \quad \text{--- } ①$$

Proof:- Since, g is a Borel Borel function on R, $g(x)$ is measurable function on ω and is a r.v. Since, g is a non-negative, its integral exists.

Hence, $E[g(x)] = \int_A g(x) + \int_{A^c} g(x)$

where $A = [|X| \geq a]$

On A, since g is non-decreasing and even (given),

$$g(a) \leq g(x) \leq a.s\ \operatorname{Sup} g(x).$$

Hence, $P(A).\, g(a) \leq \int_A g(x) \leq a.s\ \operatorname{Sup} g(x).\, P(A)$ ——— ①

On A^c, $0 \leq g(x) \leq g(a)$ and

$$0 \leq \int_{A^c} g(x) \leq g(a)\, P(A^c) \quad \text{——— } ②$$

Adding (1) & (2)

$$g(a).\,P(A) \leq E[g(x)] \leq \text{a.s } \operatorname{Sup} g(x).\, P(A) + g(a)$$

Consider,

$$E[g(x)] \leq \text{a.s } \operatorname{Sup} g(x).\, P(A) + g(a).$$

$$\Rightarrow \quad E[g(x)] - g(a) \leq \text{a.s } \operatorname{Sup} g(x).\, P(A)$$

$$\Rightarrow \quad \frac{E[g(x)] - g(a)]}{\text{a.s. } \operatorname{Sup} g(x)} \leq P(A) \quad \text{---} (3)$$

Consider, $g(a)\, P(A) \leq E[g(x)]$

$$\Rightarrow \quad P(A) \leq \frac{E[g(x)]}{g(a)} \quad \text{---} (4)$$

Combining (3) & (4)

$$\frac{E[g(x)] - g(a)}{\text{a.s } \operatorname{Sup} g(x)} \leq P(A) \leq \frac{E[g(x)]}{g(a)}$$

i.e., $$\frac{E[g(x)] - g(a)}{\text{a.s } \operatorname{Sup} g(x)} \leq P[|x| \geq a] \leq \frac{E[g(x)]}{g(a)}.$$

Hence, the Proof.

Corrolory:

If we put $g(x) = |x|^r, r > 0$ in the Basic Inequality,

We get $$\frac{E|x|^r - a^r}{\text{a.s } \operatorname{Sup} |x|^r} \leq P[|x| \geq a] \leq \frac{E[|x^r|]}{a^r}$$

which is called Markov inequality.

If we put $r = 2$ in Markov inequality, we get Chebyshev's inequality.

Convex Function:-

Definition:

Let $f(x)$ be a real valued Borel function defined on an open interval I (finite or infinite).

Then $f(x)$ is said to be convex (from below). if for every pair of points x & x' of I $\frac{x+x'}{2} \in I$

$$f\left(\frac{x+x'}{2}\right) \le \frac{1}{2} f(x) + \frac{1}{2} f(x')$$

If f is twice differentiable on I, convexity is equivalent to $f''(x) \ge$ on I.

Definition-2:

For every $x_0 \in I$, there corresponds a number $\lambda(x_0)$ such that $\forall x \in I$, $\lambda(x_0)\cdot(x-x_0) \le f(x) - f(x_0)$ — ①

The LHS of ① is the tangent through x_0 if it exists

The inequality ① implies that all the points of the curve $f(x)$ are above this tangent line

JENSON'S INEQUALITY:

If g is continuous and convex function on the interval I, and X is a r.v whose values are in I with probability 1, then $E[g(X)] \ge g[E(X)]$. provided the expectations exist.

Proof:-

To prove $E(X) \in I$:-

The various possible cases for I are:

$I = (-\infty, \infty)$; $I = (a, \infty)$; $I = (-\infty, b)$

$I = (a, b)$; $I = [a, \infty)$; $I = (-\infty, b]$. and variations of this

If $E(X)$ exists, then $-\infty < E(X) < \infty$

If $X \ge a$ (a.s), i.e., with probability 1, then $E(X) \ge a$

If $X \le b$ (a.s), ... then $E(X) \le b$

Thus $E(X) \in I$

$E(X)$ can lie i) either a left end point (if end point exists) (or) ii) a right end point of I (or) iii) an interior point of I

i) Suppose I has a left end point 'a', i.e., $X \geq a$, $E(X) \geq a$

Then $(X-a) \geq 0$ a.s and $E(X-a) = 0$.

Thus $P(X=a)=1$ (or) $P[(X-a)=0]=1$

$\therefore E[g(X)] = E[g(a)]$ $\left[\because g(X) = g(a) \text{ a.s.} \quad \because g(a) \text{ is a constant}\right.$

$= g(a)$

$= g[E(X)]$.

ii) The result can be proved similarly for right end point 'b'.

iii) Suppose $E(X) = x_0$ be an interior point of I

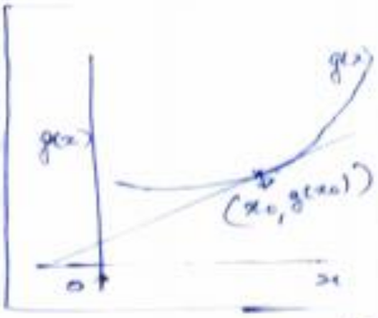
Let $ax+b$ pass through the point $(x_0, g(x_0))$ and let it be below g.

$E[g(X)] \geq E[aX+b] = aE[X]+b$

$= a x_0 + b$

$= g(x_0)$

$\Rightarrow E[g(X)] \geq g[E(X)]$



Continuous Concave Function:

A continuous function g is concave on an interval I if $(-g)$ is convex.

Corollary to the Jensen's Inequality:

If g is continuous and concave function on the interval I and X is a r.v whose values are in I with probability 1, then $E[g(X)] \leq g[E(X)]$ provided the Expectations exist.

Results:

1) If $E(X^2)$ exists then $E(X^2) \geq [E(X)]^2$

Since, $g(x) = x^2$ is convex function of x as $g''(x) = 2 > 0$.

2) If $X > 0$ a.s and $E(X)$ & $E(1/X)$ exist, then $E(1/X) \geq \frac{1}{E(X)}$ since, $g(x) = 1/x$ is a convex of x as $g''(x) = \frac{2}{x^3} > 0$ for $x > 0$.

3) If $X > 0$, (a.s), then $E[X^{1/2}] \leq [E(X)]^{1/2}$

Since $g(x) = x^{1/2}$, $x > 0$ is a concave function

as $g''(x) = -\frac{1}{4}x^{-3/2} < 0$, for $x > 0$

4) If $X > 0$, (a.s) then $E[\log(X)] \leq \log[E(X)]$,

provided the expectations exist, because

$\log x$ is a concave function of x.

5) Since, $g(x) = E(e^{tx})$ is a convex function of x $\forall t$ & $\forall x$,

If $E[e^{tX}]$ and $E(X)$ exist then

$$E[e^{tX}] \geq e^{tE(X)}$$

If $E(X) = 0$, then

$$M_X(t) = E(e^{tX}) \geq 1 \quad \forall t.$$

Thus, if $M_X(t)$ exists, then it has a lower bound 1, provided $E(X) = 0$.

This bound is attained at $t = 0$ and hence, $M_X(t)$ has a minimum at $t = 0$.

6) Let f and g be monotone functions on some subset of the real line and X be a r.v whose range is in the subset (a.s). If the expectations exist, then

$$E[f(X) \cdot g(X)] \geq E[f(X)] \cdot E[g(X)].$$

(or)

$$E[f(X) \cdot g(X)] \leq E[f(X)] \cdot E[g(X)]$$

According as f and g are monotone in the same direction

or in the opposite directions.

MARKOV'S INEQUALITY:

Let X be a r.v taking non-negative real values. Then for every $r \geq 1$,

$$P[X \geq r\,E(X)] \leq 1/r$$

i.e., If a r.v takes large values with large enough probability, then its average must also be large.

Proof:-

Let A be the event $X \geq x$.

Then $E[X] = E_A[E(X/A)]$

$$= E[X/A].P[A] + E[X/A^c].P_r[A^c]$$

Now, $E[X/A^c] \geq 0$ and $E[X/A] \geq x$,

So we get,

$$P[X \geq x] \leq \frac{E[X]}{x} \quad \text{——} (1)$$

Let $x = r.\,E(X)$.

Then (1) becomes,

$$P[X \geq r.E(X)] \leq \frac{E(X)}{r.E(X)}$$

$$\leq 1/r.$$

Hence, the Proof.

Lemma:

Let A be an event and X be a r.v.

If $P[A] \geq p+q-pq$, then $P[P(A|X) > p] \geq q$.

Proof:

Consider the compliment of A, A^c.

By assumption, $P[A^c] \leq 1-p+q+pq = (1-p)(1-q)$

Then, $E[P(A^c/X)] \leq (1-p)(1-q)$.

By Markov's Inequality,

$$P[P(A^c/X) \geq r(1-p)(1-q)] \leq 1/r.$$

Let $r = 1/(1-q)$, then $P[P(A^c/X) \geq (1-p)] \leq 1-q$

Taking the compliment of the inner event

we get, $P[P(A^c/X) < (1-p)] \geq q$.

Since, $P[A^c/X] = 1 - P[A/X]$

We get,

$$P[P[A|X) > p] \geq q.$$

When we apply this inequality, the $-pq$ term is often unimportant.

For example, Suppose we know 20% of the Indian Students do not get enough exercise. Let X be a r.v which is a random School. Let $p = q = 10\%$ which implies that in atleast 10% of the Schools, more than 10% of the Students do not get enough exercise.

End of Unit-III