# Mathematical Expectation:

The *average* value of a random phenomenon is also termed as its ***mathematical expectation*** or ***expected value***.

The expected value of a discrete r.v. is a weighted average of all possible values of the r.v., where the weights are the probabilities associated with the corresponding values.

For a discrete r.v. $\boldsymbol{X}$ with p.m.f. $\boldsymbol{p(x)}$,

$$E(X) = \sum_x x\, p(x)$$

For a continuous r.v. $\boldsymbol{X}$ with p.d.f. $\boldsymbol{f(x)}$,

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

# Expected Value of a function of a random variable

Consider a r.v. X with *p.d.f. (p.m.f.)* $f(x)$ and distribution function $F(x)$.
If $g(\cdot)$ is a function such that g(X) is a r.v. and $E[g(X)]$ exists, then

$$E[g(X)] = \sum_{x} g(x)\, p(x) \qquad \text{(for discrete r.v.)}$$

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

# Particular cases

1. If we take $g(X) = X^r$, $r$ - being a positive integer, then

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

$$E[X^r] = \sum_x x^r p(x)$$

Which is defined as $\boldsymbol{\mu_r'}$, the $r^{\text{th}}$ **moment (about origin)** of the probability distribution. Thus $\mu_r'$ ( about origin) = $E(X^r)$

In particular, if $r = 1$ $\mu_1'$ ( about origin) = $E(X)$ = $\boldsymbol{\bar{X}}$ = **Mean**

$r = 2$ $\mu_2'$ ( about origin) = $E(X^2)$ = $\int_{-\infty}^{\infty} x^2 f(x)\ dx$

$$\mu_2 = \mu_2' - {\mu_1'}^2 = E(X^2) - [E(X)]^2$$

2. If $g(X) = E[X - E(X)]^r = E[X - \bar{x}\,]^r$, then

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

$$E[X - E(X)]^r = \sum_x [x - E(X)]^r p(x) = \sum_x [x - \bar{x}\,]^r p(x)$$

which is $\boldsymbol{\mu_r}$ , the ***r*th moment about mean**.

if $r = 2$, we get $\mu_2 = E[X - E(X)]^2 = \int_{-\infty}^{\infty} [x - \bar{x}\,]^2 f(x)\ dx$ = **Variance**

3. If $g(x)$ = constant = $c$,

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

$$E(c) = \sum_x c\, p(x) = c \sum_x p(x) = c$$

# Expected Value of Two-dimensional RV

If $(X, Y)$ is a two-dimensional RV with joint *p.d.f.* $f(x, y)$ or *joint p.m.f.* $p_{ij}$, then

$$E[g(X,Y)] = \sum_i \sum_j g\left(x_i, y_j\right) p_{ij} \quad \text{(for discrete r.v.)}$$

$$E[g(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) f(x,y) \; dxdy \quad \text{(for continuous r.v.)}$$

# Properties of Expectation

1. **Addition theorem of Expectation**: $\boldsymbol{E(X+Y) = E(X) + E(Y)}$

***Proof***: $E(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)f(x,y)\ dxdy$

$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x)f(x,y)\ dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(y)f(x,y)\ dxdy$

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

***Generalisation:*** $E(X_1 + X_2 + \cdots + X_n) = E(X_1) + E(X_2) + \cdots + E(X_n)$

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

## 2. Multiplication theorem of Expectation

If $X$ and $Y$ are independent random variables, then $\boldsymbol{E(X\,Y)} = \boldsymbol{E(X) \cdot E(Y)}$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy)f(x,y)\ dxdy$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(xy)f(x)\,f(y)dxdy \qquad \text{[Since, X and Y are independent]}$$

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

$$E(X\,Y) = E(X)\,E(Y)$$

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

$$E(X_1 \cdot X_2 \cdot \cdots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \cdots \cdot E(X_n)$$

*or* $$E(\prod_{i=1}^{n} X_i) = \prod_{i=1}^{n} E(X_i)$$

3. If $\boldsymbol{X}$ is a random variable and $\boldsymbol{a}$ and $\boldsymbol{b}$ are constants, then

$$\boldsymbol{E(aX+b) = a\,E(X)+b}$$

***Proof***: $E(aX+b) = \int_{-\infty}^{\infty}(ax+b)\,f(x)\,dx$

$$= a\int_{-\infty}^{\infty} x\,f(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx$$

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

If $b=0$, then we get $E(aX) = a\,E(X)$

However,

$$E\left(\frac{1}{X}\right) \neq \frac{1}{E(X)}$$

$$E\left(X^{1/2}\right) \neq [E(X)]^{1/2}$$

$$E(\log X) \neq \log E(X)$$

$$E(X^2) \neq [E(X)]^2$$

# Variance

**Variance =** $\boldsymbol{Var}(\boldsymbol{X}) = E[X - E(X)]^2 = \mathrm{E}\left[X^2 + \left(E(X)\right)^2 - 2XE(X)\right] = E(X^2) - [E(X)]^2$

1. If $X$ is a random variable, then $V(ax + b) = a^2V(X)$

**Proof**: Let

$$Y = aX + b$$
$$\mathrm{E}(Y) = E(aX + b)$$
$$= a\,E(X) + b$$
$$Y - E(Y) = (aX + b) - (a\,E(X) + b)$$
$$= a[X - E(X)]$$
$$\mathrm{E}[Y - E(Y)]^2 = a^2\,E[X - E(X)]^2$$
$$V(Y) = a^2V(X)$$

*2.* $V(b) = 0$, *Variance of a constant is zero.*

*3.* $\sigma = \sqrt{Variance} = \sqrt{Var(X)}$ = Standard Deviation.

# Example 1

Let X be a random variable with the following probability distribution:

| | | | | |
|---|---|---|---|---|
| $X$ | : | -3 | 6 | 9 |
| $P(X = x)$ | : | 1/6 | 1/2 | 1/3 |

Find the values of $E(X)$ , $E(X^2)$ and $E(2X + 1)^2$

Solution:

$$E(X) = \sum x \cdot p(x)$$

$$= (-3) \times \frac{1}{6} + 6 \times \frac{1}{2} + 9 \times \frac{1}{3} = \frac{11}{2}$$

$$E(X^2) = \sum x^2 \cdot p(x)$$

$$= (-3^2) \times \frac{1}{6} + 6^2 \times \frac{1}{2} + 9^2 \times \frac{1}{3} = \frac{93}{2}$$

$$V(X) = E(X^2) - [E(X)]^2 = \frac{93}{2} - \left(\frac{11}{2}\right)^2 = \frac{65}{4}$$

$$E(2X+1)^2 = E(4X^2 + 4X + 1) = 4\,E(X^2) + 4\,E(X) + 1$$

$$= 4 \times \frac{93}{2} + 4 \times \frac{11}{2} + 1 = 209$$

# Example 2

Find the expectation of the number on a die when thrown.

**Solution:**

Let $X$ be the random variable representing the number on a die when thrown. Then $X$ can take any one of the values $1, 2, 3, 4, 5, 6$ each with equal probability ${}^{1}/{}_{6}$.

Hence, $$E(X) = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \frac{1}{6} \times 3 + \cdots + \frac{1}{6} \times 6$$

$$= \frac{1}{6}\,(1 + 2 + 3 + \cdots + 6) = \frac{1}{6} \times \frac{6 \times 7}{2} = \frac{7}{2}$$

# Example 3:

Two unbiased dice are thrown. Find the expected values of the sum of numbers of points on them.

# Co-Variance

If X and Y are two random variables, then the covariance between them is defined as:

$$Cov(X,Y) = E[\{X - E(X)\}\{Y - E(Y)\}]$$

$$= E[XY - X\,E(Y) - Y\,E(X) + E(X)E(Y)]$$

$$= \mathrm{E}(XY) - E(X)\,E(Y) - E(Y)\,E(X) + E(X)E(Y)$$

$$Cov(X,Y) = \mathrm{E}(XY) - E(X)\,E(Y)$$

If $X$ and $Y$ are independent random variables, $\mathrm{E}(XY) = E(X)\,E(Y)$

and hence, $Cov(X,Y) = E(X)\,E(Y) - E(X)\,E(Y) = 0$

# Properties

1. $Cov(aX, bY) = E[\{aX - E(aX)\}\{bY - E(bY)\}]$

$= E[a\{X - E(X)\}\, b\{Y - E(Y)\}]$

$= ab\, E[\{X - E(X)\}\, \{Y - E(Y)\}]$

$Cov(X, Y) \quad = ab\, Cov(X, Y)$

2. $Cov(X + a, Y + b) = Cov(X, Y)$

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

4. If X and Y are independent, $Cov(X, Y) = 0$, but the converse is not true.

5. $V(X_1 \pm X_2) = V(X_1) + \mathrm{V}(X_2) \pm 2Cov(X_1, X_2)$

6. $V(\sum_{i=1}^{n} a_i X_i) = \sum_{i=1}^{n} a_i^2\, V(X_i) + 2\, \sum_{i=1}^{n} \sum_{j=1}^{n} a_i a_j\, Cov(X_i, X_j)$

# Correlation Coefficient

If $X$ and $Y$ are two random variables, then the covariance between them is denoted as $Cov(X, Y)$,

$$Cov(X, Y) \; = \; \mathrm{E}(XY) - E(X)\,E(Y)$$

and the Co-efficient of correlation($r_{XY}$ $or$ $\rho_{XY}$) between X and Y is defined as ***a numerical measure of linear relationship between them***,

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(x)}\,\sqrt{Var(y)}}$$

# Properties

1. If $X$ and $Y$ are independent, $Cov(X,Y) = 0$ and $\rho_{XY} = 0$

   but the converse is not true.

   When $\rho_{XY} = 0$, we say that X and Y are uncorrelated.

2. $|\rho_{XY}| \leq 1$

   i.e., if $\rho_{XY} = +1 \Rightarrow$ Positive Correlation

   $\rho_{XY} = -1 \Rightarrow$ Negative Correlation

   $\rho_{XY} = 0 \Rightarrow$ UnCorrelated random variables.

# Skewness

When a series is not symmetrical it is said to be asymmetrical or *skewed*.

A distribution is said to be ‘skewed’ when the ***mean*** and ***median*** fall at different points in the distribution, and the balance (or centre of gravity) is shifted to one side or the other to left or right.

***Measure of skewness is the lack of symmetry of a distribution.***

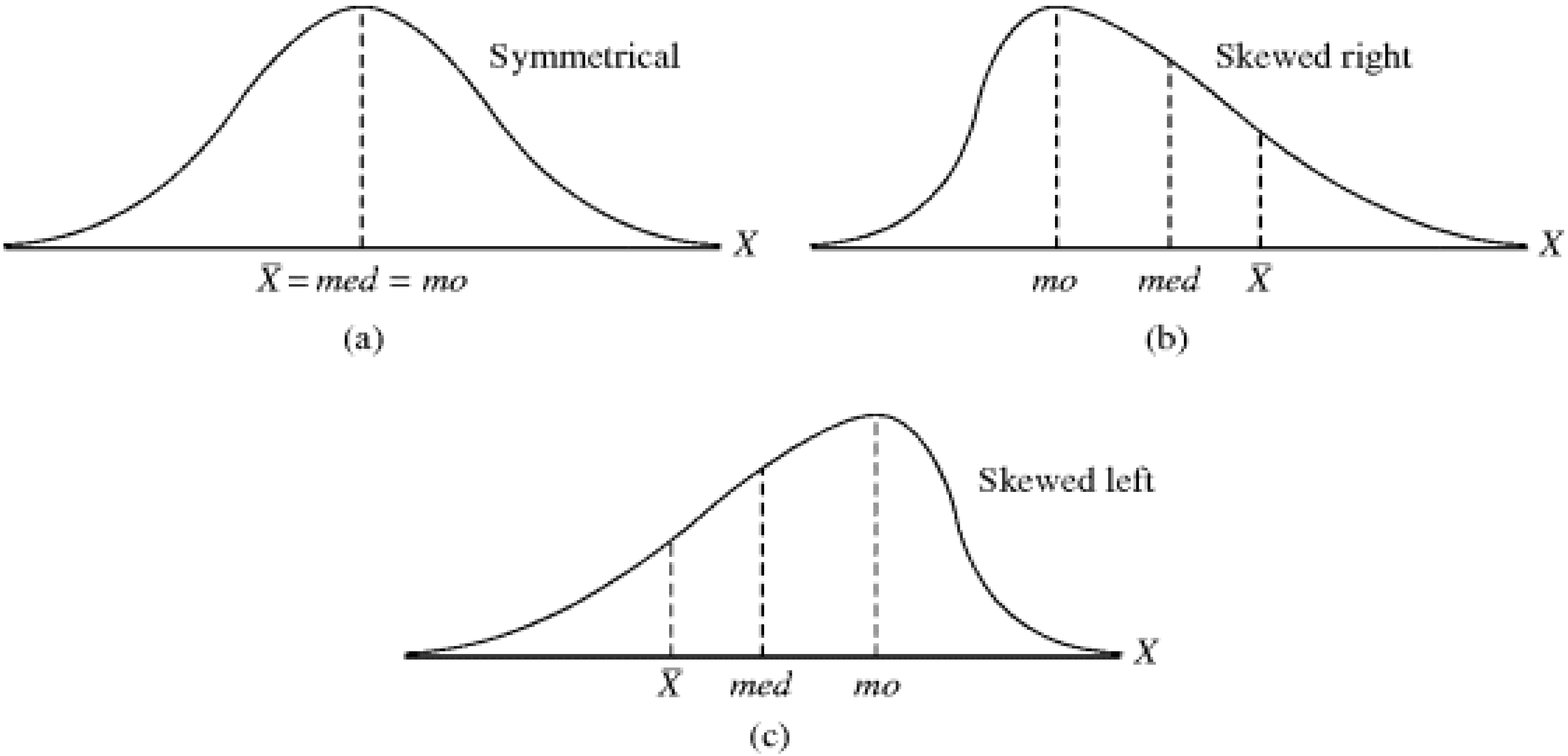
Dispersion is concerned with the amount of variation rather than with its direction.

Skewness tell us about the ***direction of the variation*** or the departure from symmetry.

Types of Skewness: (i) Symmetrical Distribution

(ii) Positively Skewed Distribution (**Mean > Mode**)

(iii) Negatively Skewed Distribution (**Mean < Mode**)


Symmetrical
X
X̄ = med = mo
(a)
Skewed right
X
mo
med
X̄
(b)
Skewed left
X
X̄
med
mo
(c)

# (iii) Moments

Moment is a measure of a force with respect to its tendency to provide rotation.

The strength of the tendency depends on the amount of force and the distance from the origin of the point at which the force is exerted.

*Definition*: Let the symbol $\boldsymbol{x_i = (X_i - \bar{X})}$ be used to represent the deviation of any item in a distribution from the arithmetic average of that distribution.

The arithmetic mean of the various powers of these deviations in any distribution are called the ***moments*** of the distribution.

# **Central Moments** (Moments about the Arithmetic Mean):

First Moment $\mu_1 = \frac{\sum(X_i - \bar{X})}{N}$ (sum of the deviations from A.M. is always zero. $\boldsymbol{\mu_1 = 0}$)

Second Moment $\mu_2 = \frac{\sum(X_i - \bar{X})^2}{N} = \sigma^2$ = Variance

Third Moment $\mu_3 = \frac{\sum(X_i - \bar{X})^3}{N}$

Fourth Moment $\mu_4 = \frac{\sum(X_i - \bar{X})^4}{N}$

# **Non-central Moments** (Moments about the Assumed Mean)

Where the actual mean is in fractions it is difficult to calculate moments by applying the above formulae. In such cases we can first compute moments about an **arbitrary origin($A$)**.

$$\mu_1' = \frac{\sum(X_i - A)}{N} = \bar{X} - A \quad ; \quad \mu_1' = \frac{\sum f_i(X_i - A)}{N}$$

$$\text{Mean} = \bar{X} = \mu_1' + \text{A}$$

$$\mu_2' = \frac{\sum(X_i - A)^2}{N} \quad ; \quad \mu_2' = \frac{\sum f_i(X_i - A)^2}{N}$$

$$\mu_3' = \frac{\sum(X_i - A)^3}{N} \quad ; \quad \mu_3' = \frac{\sum f_i(X_i - A)^3}{N}$$

$$\mu_4' = \frac{\sum(X_i - A)^4}{N} \quad ; \quad \mu_4' = \frac{\sum f_i(X_i - A)^3}{N}$$

# Conversion of moments about an ***Arbitrary origin*** into Moments about ***mean***

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

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

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

$$\mu_4 = \mu_4' - 4\mu_1'\mu_3' + 6\mu_2'(\mu_1')^2 - 3(\mu_1')^4$$

$\beta_1$ (beta one) = $\frac{\mu_3^2}{\mu_2^3}$ (Coeff. of Skewness)

$\beta_2$ (beta two) = $\frac{\mu_4}{\mu_2^2}$ (Coeff. of Kurtosis)

$\gamma_1$(Gamma one) = $\sqrt{\beta_1}$ (Coeff. of Skewness)

$\gamma_2$ (Gamma two) = $\beta_2 - 3$ (Coeff. of Kurtosis)

# Kurtosis

Kurtosis enables us to have an idea about the '*flatness*' or '*peakedness*' in the region about the mode of a frequency curve.

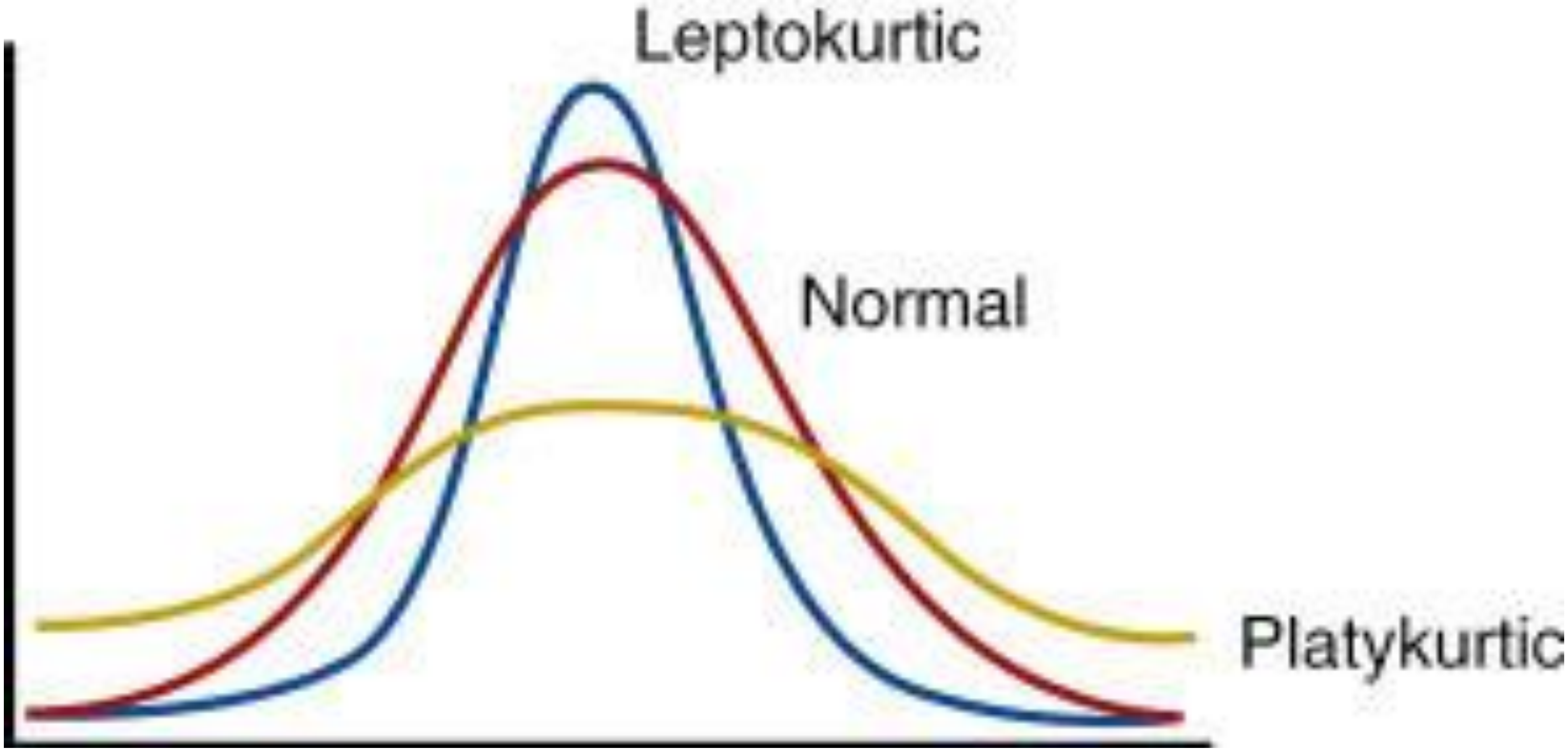
*"Convexity of the frequency curve"*

It is measured by the Coefficient of $\beta_2 = \frac{\mu_4}{\mu_2^2}$ or $\gamma_2 = \beta_2 - 3$

(i) $\beta_2 = 3$, *i.e.*, $\gamma_2 = 0$ : mesokurtic curve

(ii) $\beta_2 < 3$, *i.e.*, $\gamma_2 < 0$ : platykurtic curve

(iii) $\beta_2 > 3$, *i.e.*, $\gamma_2 > 0$ : leptokurtic curve


Leptokurtic
Normal
Platykurtic

**Example 4**: For a bivariate probability distribution of $(X, Y)$ given below,

| **Y** \ **X** | **-1** | **0** | **1** |
|---|---|---|---|
| **-1** | 0 | 0.1 | 0.1 |
| **0** | 0.2 | 0.2 | 0.2 |
| **1** | 0 | 0.1 | 0.1 |

Find (i) Prove that $X$ and $Y$ are uncorrelated.

(ii) Find $Var\,(X)$ and $Var(Y)$ .

(iii) Find $r(X, Y)$.

Solution

(i) $E(Y) = \sum_{y_j} y_i P\left(Y = y_j\right) \; = \; -1 \times (0.2) + 0 \times (0.6) + 1 \times (0.2) = 0$

$E(X) = \sum x_i \, P(X = x_i) \; = \; -1 \times (0.2) + 0 \times (0.4) + 1 \times (0.4) = 0.2$

(ii) 
$$
\begin{aligned}
E(XY) &= \sum x_i \, y_j \, p_{ij} \\
&= (-1)(-1)(0) + (0)(-1)(0.1) + 1(-1)(0.1) \\
&\quad + (0)(-1)(0.2) + (0)(0)(0.2) + (0)(1)(0.2) \\
&\quad +1(-1)(0) + 1(0)(0.1) + 1(1)(0.1)
\end{aligned}
$$

$E(XY) \quad = \; -0.1 + 0.1 = 0$

$\therefore \qquad Cov(X,Y) = E(XY) - E(X)E(Y) = 0$

$\Longrightarrow \qquad X$ and $Y$ are uncorrelated.

(iii) $E(X^2) = \sum x^2\, P(X = x_i) \;=\; (-1)^2(0.2) + 0(0.4) + 1^2(0.4) = 0.6$

$$V(X) = E(X^2) - \big(E(X)\big)^2 = 0.56$$

$$E(Y^2) = \sum y^2\, P(Y = y_i) \;=\; (-1)^2(0.2) + 0(0.6) + 1^2(0.2) = 0.4$$

$$V(Y) = E(Y^2) - \big(E(Y)\big)^2 = 0.4$$

$$\mathrm{r} = \rho_{XY} = \frac{Cov(X,Y)}{\sigma_X\,\sigma_Y} = \frac{Cov(X,Y)}{\sqrt{Var(x)}\,\sqrt{Var(y)}} = 0$$

# Example 5

Two random variables X and Y have the following joint probability density function:

$$f(x,y) = \begin{cases} 2 - x - y\,; & 0 \le x \le 1\ , 0 \le y \le 1 \\ 0\ ; & otherwise \end{cases}$$

Find (i) Marginal probability density function of X and Y.

(ii) Var (X), Var (Y) and $Cor(X,Y)$.

Solution

Given, $$f(x,y) = \begin{cases} 2-x-y\,;\; 0 \le x \le 1\,, 0 \le y \le 1 \\ 0\,; \quad otherwise \end{cases}$$

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

$$= \int_0^1 (2-x-y)\,dy = \frac{3}{2} - x$$

$$\therefore \qquad f(x) = \frac{3}{2} - x\,, \qquad 0 < x < 1$$

Similarly, $$f(y) = \frac{3}{2} - y\,, \qquad 0 < y < 1$$

(ii) $E(X) = \int_{-\infty}^{\infty} x\, f(x)\, dx = \int_0^1 x \left(\frac{3}{2} - x\right) dx = \frac{5}{12}$

$$E(Y) = \int_{-\infty}^{\infty} y\, f(y)\, dy = \int_0^1 y \left(\frac{3}{2} - y\right) dy = \frac{5}{12}$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2\, f(x)\, dx = \int_0^1 x^2 \left(\frac{3}{2} - x\right) dx = \frac{1}{4}$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2\, f(y)\, dy = \int_0^1 y^2 \left(\frac{3}{2} - y\right) dy = \frac{1}{4}$$

$$V(X) = E(X^2) - \left(E(X)\right)^2 = \frac{11}{144} = V(Y)$$

$$E(XY) = \int_0^1 \int_0^1 xy \ (2 - x - y) \, dx \, dy$$

$$E(XY) = \frac{1}{6}$$

$$\therefore \qquad Cov(X, Y) = E(XY) - E(X)E(Y) = -\frac{1}{144}$$

# Conditional Expectation

If $(X, Y)$ is a two-dimensional discrete random variable with joint probability mass function $P\left(X = x_i, Y = y_j\right)$, then the conditional expectation of $g(X, Y)$ are defined as:

$$E\left\{g(X,Y)/\ Y = y_j\right\} \ = \ \sum_i \ g\left(x_i, y_j\right) \times P\left(X = x_i\ /\ Y = y_j\right)$$

$$= \ \sum_i \ g\left(x_i, y_j\right) \times \frac{P\left(X=x_i, Y=y_j\right)}{P\left(Y=y_j\right)}$$

and $\quad E\{g(X,Y)/\ X = x_i\} \ = \ \sum_j \ g\left(x_i, y_j\right) \times \frac{P\left(X=x_i, Y=y_j\right)}{P(X=x_i)}$

If $(X, Y)$ is a two-dimensional continuous random variable with joint probability density function $f(x, y)$, then

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

$$= \int_{-\infty}^{\infty} g(x, y) \times \frac{f(x,y)}{f(y)}\, dx$$

and $$E\{g(X,Y)/\, X = x_i\} = \int_{-\infty}^{\infty} g(x, y) \times \frac{f(x,y)}{f(x)}\, dy$$

# Conditional Mean and Variances

Conditional Mean:

$$\mu_{X/Y} = E(X/Y) \quad = \int_{-\infty}^{\infty} x\, f(x/y)\, dx$$

$$\mu_{Y/X} = E(Y/X) \quad = \int_{-\infty}^{\infty} y\, f(y/x)\, dy$$

Conditional Variances:

$$\sigma_{X/Y}^2 = E\left\{\left(X - \mu_{X/Y}\right)^2\right\} = \int_{-\infty}^{\infty} \left(X - \mu_{X/Y}\right)^2 f(x/y)\, dx$$

$$\sigma_{Y/X}^2 = E\left\{\left(Y - \mu_{Y/X}\right)^2\right\} = \int_{-\infty}^{\infty} \left(Y - \mu_{Y/X}\right)^2 f(y/x)\, dy$$

**Example 6**: Let $X$ and $Y$ be a two random variables each taking three values -1, 0 and 1 having the joint probability distribution

| Y \ X | **-1** | **0** | **1** |
|---|---|---|---|
| **-1** | 0 | 0.1 | 0.1 |
| **0** | 0.2 | 0.2 | 0.2 |
| **1** | 0 | 0.1 | 0.1 |

Find $Var\ (Y \mid X = -1)$ and $Var\ (X \mid Y = -1)$

# Example 7

Let $(X, Y)$ be a two random variable having the joint probability density function:

$$f(x, y) = \begin{cases} 8xy & ; \quad 0 < x < y < 1 \\ 0 & ; \quad elsewhere \end{cases}$$

Find (i) $E(Y \mid X = x)$ and $E(X \mid Y = y)$

(ii) $Var\,(Y \mid X = x)$ and $Var\,(X \mid Y = y)$

# Example 8

Let $(X, Y)$ be a two random variable having the joint probability density function:

$$f(x,y) = \begin{cases} xe^{-x(y+1)} & ; \quad x \geq 0\,, y \geq 0 \\ \quad 0 \quad ; & \quad elsewhere \end{cases}$$

Find the regression curve of $\boldsymbol{Y}$ **on** $\boldsymbol{X}$.

# Bounds on Probability

## Markov's Inequality

If $\boldsymbol{X}$ is a random variable that takes only non-negative values, then for any value $a > 0$,

$$P\{X \geq a\} \leq \frac{E[X]}{a}$$

As a corollary, we obtain **Chebyshev's inequality:**

If $\boldsymbol{X}$ is a random variable with mean $\boldsymbol{\mu}$ and variance $\boldsymbol{\sigma^2}$, then for any value $k > 0$,

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2}$$

Note: The importance of Markov's and Cheybyshev's inequalities is that they enable us to derive bounds on probabilities when only the **mean**, or both the **mean and the variance**, of the probability distribution are known. Of course, if the actual distribution were known, then the desired probabilities could be exactly computed and we would not need to resort to bounds.

# Example 9

Suppose that it is known that the number of items produced in a factory during a week is a random variable with mean 50,

(a) What can be said about the probability that this week's production will exceed 75?

(b) If the variance of a week's production is known to equal 25, then what can be said about the probability that this week's production will be between 40 and 60?

# Example 10

If $X$ denotes the sum of the numbers obtained when 2 dice are thrown, obtain an upper bound for $P\{|X - 7| \geq 4\}$. Compare with the exact probabilities.

**Cauchy-Schwartz Inequality**

$$[E(XY)]^2 \leq E(X^2) \cdot E(Y^2)$$

**Jenson's Inequality**

If $\boldsymbol{g}$ is a continuous and convex function on the interval I and X is a random variable whose values are in $\boldsymbol{I}$ with probability 1, then

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

provided the expectations exist.

# Weak law of large numbers

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables, each having mean $E[X_i] = \mu$. Then, for any $\varepsilon > 0$,

$$P\left\{\left|\frac{X_1+X_2+\ldots+X_n}{n} - \mu\right| > \varepsilon\right\} \to 0 \quad \text{as } n \to \infty$$

We shall prove the result only under the additional assumption that the random variables have a finite variance $\sigma^2$. Now, as

$$E\left[\frac{X_1 + \cdots + X_n}{n}\right] = \mu \quad \text{and} \quad \operatorname{Var}\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{\sigma^2}{n}$$

it follows from Chebyshev's inequality that

$$P\left\{\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right\} \leq \frac{\sigma^2}{n\epsilon^2}$$

# Central limit theorem

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables, each having mean $E[X_i] = \mu$ and variance $\mathrm{V}(X_i) = \sigma^2$, $\mathrm{i} = 1, 2, \ldots$

Then for $\boldsymbol{n}$ is large($n \to \infty$), the distribution of $X_1 + X_2 + \ldots + X_n$

Is approximately normally distributed with mean $\boldsymbol{nu}$ and variance $\boldsymbol{n\sigma^2}$.

# Example 11

An insurance company has 25,000 automobile policy holders. If the yearly claim of a policy holder is a random variable with mean 320 and standard deviation 540, approximate the probability that the total yearly claim exceeds 8.3 million.