Statistical Averages

UNIT – 2

Instructor: Dr. Sahaya Jenifer A

Expectation

- a "representative" value of an event X.

- a weighted average of the possible values of X.

Expectation – a simple game

I flip a coin. If the head comes up, you pay me Rs. 5. If the tail comes up, I pay you Rs. 3. Now, imagine we play this many times, say 100. Then, what is ***my expected income*** for a single game?

The coin has $P(H) = 0.5$ and $P(T) = 0.5$. So, we *expect* to see about $0.5 \times 100 = 50$ heads and $0.5 \times 100 = 50$ tails.

For 100 coin flips, my total income should be

$$50 \times 5 - 50 \times 3 = \text{Rs.}\,100.$$

For a single coin flip, **my expected income** should be

$$0.5 \times 5 - 0.5 \times 3 = \text{Rs.}\,1$$

Mathematical Expectation (the mean)

If X is a random variable which can assume

any one of the values $x_1, x_2, \ldots, x_n$ with

respective probabilities $p_1, p_2, \ldots, p_n$, then

$$E(X) = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n = \sum x_i p_i$$

Mathematical Expectation (the mean)

If X is a random variable which can assume any one of the values in $[-\infty, \infty]$ with probability density function $f(x)$, then

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

Properties of Expectation

(1) $E(c) = c$, where c is a constant

(2) $E(cX) = c\, E(X)$, where c is a constant

(3) $E(aX + b) = a\, E(X) + b$, where a and b are constants.

(4) If X and Y are random variables, then

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

(5) If X and Y are ***independent*** random variables, then

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

What is the value of $E[E[x] + 1]$?

$a)$ $E[X] + 1$

$b)$ 1

$c)$ 0

Example 1: What is the expected number of heads appearing when a fair coin is tossed three times?

Solution

$\boldsymbol{x}$	0	1	2	3
$\boldsymbol{p(x)}$	1/8	3/8	3/8	1/8

$$E[X] = \sum k_i p_i$$
$$= 0\left(\frac{1}{8}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{3}{8}\right) + 3\left(\frac{1}{8}\right)$$
$$= \frac{3}{8} + \frac{6}{8} + \frac{3}{8} = \frac{12}{8} = \frac{3}{2}$$

Example 2: Consider the PMF given by

$$p_X(k) = \begin{cases} 1/16, & \text{if } k = 0 \\ 3/8, & \text{if } k = 1 \\ 9/16, & \text{if } k = 2 \end{cases}$$

Calculate the mean/expectation.

Solution

$$p_X(k) = \begin{cases} 1/16, & \text{if } k = 0 \\ 3/8, & \text{if } k = 1 \\ 9/16, & \text{if } k = 2 \end{cases}$$

$$E[X] = \sum k_i p_i$$

$$= 0\left(\frac{1}{16}\right) + 1\left(\frac{3}{8}\right) + 2\left(\frac{9}{16}\right)$$

$$= \frac{3}{8} + \frac{9}{8} = \frac{12}{8} = \frac{3}{2}$$

Variance of $\boldsymbol{X}$

- a measure of dispersion of X around its mean.

$$\begin{aligned} \mathrm{Var}(X) = \sigma^2 &= E[X - E(X)]^2 \\ &= E(X^2) - [E(X)]^2 \\ &= \Sigma x_i^2 p_i - [\Sigma\, x_i p_i]^2 \end{aligned}$$

Properties of Variance

(i) $Var(c) = 0$, where c is a constant

(ii) $Var(X \pm c) = Var(X)$

(iii) $Var(aX) = a^2$, where a is a constant

(iv) $Var(aX \pm b) = Var(aX) = a^2\, Var(X)$

Note

- Variance is always nonnegative.

- Standard deviation is practically useful,

because it has the same units as X.

Example 3: The number of hardware failures of a computer system in a week of operations has the following pmf

No. of failures	0	1	2	3	4	5	6
Probability	0.18	0.28	0.25	0.18	0.06	0.04	0.01

(i) Find the mean of the number of failures in a week

(ii) Find the variance.

Solution

$$E(X) = \sum x\, P(X = x)$$

$$= 0(0.18) + 1(0.28) + 2(0.25) + 3(0.18) + 4(0.06) + 5(0.04) + 6(0.01)$$

$$= 0 + 0.28 + 0.50 + 0.54 + 0.24 + 0.20 + 0.06$$

$$= 1.82$$

$$E(X^2) = \sum x^2 \, P(X = x)$$

$$= 0^2(0.18) + 1^2(0.28) + 2^2(0.25) +$$
$$3^2(0.18) + 4^2(0.06) + 5^2(0.04) + 6^2(0.01)$$

$$= 0 + 0.28 + 1 + 1.62 + 0.96 + 1 + 0.36$$

$$= 5.22$$

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

$$= 5.22 - (1.82)^2 = 1.9076$$

Example 4

When a single fair die is thrown, X denotes the number that turns up. Find $E(X), E(X^2)$ and $Var(X)$.

Solution

For a fair die, X can be any of the value in $\{1,2,3,4,5,6\}$ with equal probability 1/6.

$$E(X) = \sum_{x=1}^{6} x\, P(X = x)$$

$$= 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right)$$

$$+ 6\left(\frac{1}{6}\right) = 21\left(\frac{1}{6}\right) = 3.5$$

$$E(X^2) = \sum_{x=1}^{6} x^2 P(X = x)$$

$$= 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{1}{6}\right) + 3^2\left(\frac{1}{6}\right) + 4^2\left(\frac{1}{6}\right)$$

$$+ 5^2\left(\frac{1}{6}\right) + 6^2\left(\frac{1}{6}\right) = 91\left(\frac{1}{6}\right) = 15.1667$$

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

$$= 15.1667 - 3.5^2 = 2.9167$$

Example 5

Consider the following probability function

$X = x_i$	-2	-1	0	1	2	3
$P(X = x_i)$	0.1	k	0.2	$2k$	0.3	$3k$

Find (i) the value of k

(ii) $P(X < 2)$ and $P(-2 < x < 2)$

(iii) the cdf of X

(iv) the mean and variance of X

Solution

(i) The sum of all probabilities must be equal to 1.

That is, $\sum P(X = x_i) = 1$.

$$0.1 + k + 0.2 + 2k + 0.3 + 3k = 1$$

$$6k + 0.6 = 1$$

$$6k = 1 - 0.6 = 0.4 \Rightarrow k = \frac{0.4}{6} = \frac{1}{15}$$

(ii)

$$
\begin{aligned}
P(X < 2) = P(X = -2) + P(X = -1) \\
+P(X = 0) + P(X = 1) \\
= 0.1 + k + 0.2 + 2k \\
= 0.1 + \frac{1}{15} + 0.2 + \frac{2}{15} = 0.5
\end{aligned}
$$

$$P(-2 < X < 2) = P(X = -1) + P(X = 0)$$
$$+P(X = 1)$$
$$= k + 0.2 + 2k$$
$$= \frac{1}{15} + 0.2 + \frac{2}{15} = 0.4$$

(iii) CDF $F(x) = P(X \leq x)$

$$F(-2) = P(X \leq -2) = P(X = 2) = 0.1$$

$$F(-1) = P(X \leq -1) = 0.1 + \frac{1}{15} = 0.1667$$

$$F(0) = P(X \leq 0) = 0.1667 + 0.2 = 0.3667$$

$$F(1) = P(X \leq 1) = 0.3667 + \frac{2}{15} = 0.5$$

$$F(2) = P(X \leq 2) = 0.5 + 0.3 = 0.8$$

$$F(3) = P(X \leq 3) = 0.8 + \frac{3}{15} = 1$$

The CDF is

$$F(x) = \begin{cases} 0, & x < -2 \\ 0.1, & -2 \leq x < -1 \\ 0.1667, & -1 \leq x < 0 \\ 0.3667, & 0 \leq x < 1 \\ 0.5, & 1 \leq x < 2 \\ 0.8, & 2 \leq x < 3 \\ 1, & x \geq 3 \end{cases}$$

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

$$= -2(0.1) + (-1)\left(\frac{1}{15}\right) + 0(0.2) + 1\left(\frac{2}{15}\right)$$

$$+ 2(0.3) + 3\left(\frac{3}{15}\right)$$

$$= 1.0667$$

$$E(X^2) = \sum x_i^2 P(X = x_i)$$

$$= (-2)^2(0.1) + (-1)^2\left(\frac{1}{15}\right) + 0^2(0.2)$$

$$+ 1^2\left(\frac{2}{15}\right) + 2^2(0.3) + 3^2\left(\frac{3}{15}\right)$$

$$= 4(0.1) + \left(\frac{1}{15}\right) + \left(\frac{2}{15}\right) + 4(0.3)$$

$$+ 9\left(\frac{3}{15}\right) = 3.6$$

$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= 3.6 - (1.07)^2 \\
&= 2.455
\end{aligned}
$$

Example 6

If the pmf of a RV X is given by

$$P(X = r) = kr^3, \qquad r = 1,2,3,4$$

Find (i) the value of k

(ii) the mean, variance of X

(iii) the cdf of X

(iv) $P\left[\frac{1}{2} < X < \frac{5}{2} \middle| X > 1\right]$

Solution

(i) The sum of all probabilities must equal 1.

That is, $\sum_{r=1}^{4} P(X = r) = 1.$

$$k(1)^3 + k(2)^3 + k(3)^3 + k(4)^3 = 1$$

$$k + 8k + 27k + 64k = 1$$

$$100k = 1 \Rightarrow k = \frac{1}{100} = 0.01$$

(ii)

$$
\begin{aligned}
E(X) &= \sum rP(X = r) \\
&= 1(1^3k) + 2(2^3k) + 3(3^3k) + 4(4^3k) \\
&= k + 16k + 81k + 256k \\
&= 354k = \frac{354}{100} = 3.54
\end{aligned}
$$

$$
\begin{aligned}
E(X^2) &= \sum_{r=1}^{4} r^2 P(X=r) \\
&= 1^2(1^3k) + 2^2(2^3k) + 3^2(3^3k) + 4^2(4^3k) \\
&= k + 32k + 243k + 1024k \\
&= 1300k = \frac{1300}{100} = 13
\end{aligned}
$$

$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= 13 - (3.54)^2 = 0.4684
\end{aligned}
$$

(iii) CDF $F(x) = P(X \leq x)$

$$
\begin{aligned}
F(1) &= P(X = 1) = 1^3k = k = 0.01 \\
F(2) &= P(X \leq 2) = 0.01 + 2^3k \\
&= 0.01 + 0.08 = 0.09 \\
F(3) &= P(X \leq 3) = 0.09 + 3^3k \\
&= 0.09 + 0.27 = 0.36 \\
F(4) &= P(X \leq 4) = 0.36 + 4^3k \\
&= 0.36 + 0.64 = 1
\end{aligned}
$$

The CDF is

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.01, & 1 \leq x < 2 \\ 0.09, & 2 \leq x < 3 \\ 0.36, & 3 \leq x < 4 \\ 1, & x \geq 4 \end{cases}$$

(iv) Conditional probability: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Here $A = \left\{\frac{1}{2} < X < \frac{5}{2}\right\} = \{X = 1, X = 2\}$

$B = \{X > 1\} = \{X = 2, X = 3, X = 4\}$

$P(A \cap B) = P(X = 2) = 2^3 k = 0.08$

$P(B) = P(X = 2) + P(X = 3) + P(X = 4)$

$= 0.99$

$$P\left(\frac{1}{2} < X < \frac{5}{2} \middle| X > 1\right) = \frac{0.08}{0.99} = 0.0808$$

Example 7

Find the mean and S.D. of the following pdf

$$f(x) = \begin{cases} kx(2-x), & 0 \le x \le 2 \\ 0, & \text{elsewhere} \end{cases}$$

To find k:

The PDF satisfy the condition $\int_{-\infty}^{\infty} f(x)\,dx = 1$

$$\int_0^2 kx(2-x) = 1$$

$$k\int_0^2 (2x - x^2) = 1$$

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

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

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

$$k\left(4-\frac{8}{3}\right)=1$$

$$k\left(\frac{4}{3}\right)=1$$

$$k=\frac{3}{4}$$

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

$$= \int_0^2 x \frac{3}{4} x(2 - x) dx$$

$$= \int_0^2 \frac{3}{4} x^2 (2 - x) dx$$

$$= \frac{3}{4} \int_0^2 (2x^2 - x^3) dx$$

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

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

$$= \frac{3}{4}\left[\frac{16}{3} - 4\right]$$

$$= \frac{3}{4}\left[\frac{4}{3}\right] = 1$$

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

$$= \int_0^2 x^2 \frac{3}{4} x(2 - x)dx$$

$$= \frac{3}{4}\int_0^2 (2x^3 - x^4)\, dx$$

$$= \frac{3}{4}\left[\frac{2x^4}{4} - \frac{x^5}{5}\right]_0^2 = \frac{3}{4}\left[\frac{x^4}{2} - \frac{x^5}{5}\right]_0^2$$

$$= \frac{3}{4}\left[\frac{2^4}{2} - \frac{2^5}{5}\right]$$

$$= \frac{3}{4}\left[\frac{16}{2} - \frac{32}{5}\right]$$

$$= \frac{3}{4}\left[8 - \frac{32}{5}\right] = \frac{3}{4}\left[\frac{8}{5}\right] = \frac{6}{5}$$

$$E(X^2) = \frac{6}{5}$$

Example 8

A continuous random variable has pdf

$$f(x) = \begin{cases} a + bx, & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$$

If the mean of the distribution is 1/2, then

find a and b. Also, compute $Var(X)$.

Solution

The PDF should satisfy the condition

$$\int_0^1 f(x)dx = 1$$

$$\int_0^1 (a + bx)dx = 1$$

$$\left[ax + \frac{bx^2}{2}\right]_0^1 = 1$$

$$a + \frac{b}{2} = 1$$

$$2a + b = 2 ------- (1)$$

Given that $E(X) = \frac{1}{2} \Rightarrow \int_0^1 x\, f(x) dx = \frac{1}{2}$

$$\int_0^1 x(a + bx) dx = \frac{1}{2}$$

$$\int_0^1 (ax + bx^2) dx = \frac{1}{2}$$

$$\left[\frac{ax^2}{2} + \frac{bx^3}{3}\right]_0^1 = \frac{1}{2}$$

$$\frac{a}{2} + \frac{b}{3} = \frac{1}{2}$$

$$3a + 2b = 3 -------- -(2)$$

Solving (1) and (2),

(1)x2-> $4a + 2b = 4$

(2)-> $3a + 2b = 3 \; (-)$

$a \qquad\qquad = 1$

From (1), $b = 2 - 2a = 2 - 2(1) = 0$

So, $f(x) = 1 + 0x = 1$

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

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

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

$$= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{12}$$

Example 9

A test engineer proposed, after a series of tasks, that the life time X of a component is a random variable with pdf $f(x) = \begin{cases} \frac{x}{180} e^{-x/10}, & x \geq 0 \\ 0, & \text{otherwise} \end{cases}$

Find (i) $E(X)$ (ii) $Var(X)$

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

$$= \int_0^\infty x \frac{x}{180} e^{-x/10} dx$$

$$= \frac{1}{180} \int_0^\infty x^2\, e^{-x/10} dx$$

Gamma function: $\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}}$ for $a > 0$

Here, $n = 2, a = 1/10 = 0.1$

$$E(X) = \frac{1}{180} \frac{2!}{(0.1)^3} = 11.11$$

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

$$= \int_0^\infty x^2 \frac{x}{180} e^{-x/10} dx$$

$$= \frac{1}{180} \int_0^\infty x^3\, e^{-x/10} dx$$

Gamma function: $\int_0^\infty x^n e^{-ax} dx = \frac{n!}{a^{n+1}}$ for $a > 0$

Here, $n = 3, a = 1/10 = 0.1$

$$E(X) = \frac{1}{180} \frac{3!}{(0.1)^4} = 333.33$$

$$
\begin{aligned}
Var(X) &= E(X^2) - [E(X)]^2 \\
&= 333.33 - (11.11)^2 \\
&= 209.90
\end{aligned}
$$

Example 7

The cdf of a random variable X is $f(x) = 1 - (1 + x)e^{-x},\ x > 0$. Find the pdf of X; mean and variance of X.

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

$$f(x) = \frac{d}{dx}[1 - (1 + x)e^{-x}]$$

$$= -\frac{d}{dx}[(1 + x)e^{-x}]$$

$$= -[(1 + x)(-e^{-x}) + e^{-x}(1)]$$

$$= -[-e^{-x} - xe^{-x} + e^{-x}]$$

$$= -[-xe^{-x}] = xe^{-x}$$

$$f(x) = xe^{-x}, x > 0$$

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

$$= \int_0^\infty x(xe^{-x}) dx = \int_0^\infty x^2 e^{-x} dx$$

Gamma function: $\int_0^\infty x^n e^{-x} dx = n!$

Here, $\mathrm{n} = 2$

$$E(X) = 2! = 2$$

(iii) $E(X^2) = \int_0^\infty x^2 \, f(x)dx$

$$= \int_0^\infty x^2(xe^{-x})dx = \int_0^\infty x^3 e^{-x} dx$$

Gamma function: $\int_0^\infty x^n e^{-x} dx = n!$

Here, $\mathrm{n} = 3$

$$E(X^2) = 3! = 6$$

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

$$= 6 - 2^2 = 2$$

Moments

- The expected value of an integral power of a random variable is called its moment.
- Helps to find the central tendency, dispersion, skewness and the peakedness of the curve.

Moments about the mean (Central moments)

The r −th moment of a RV X about the mean μ is given by

$$\mu_r = E[(X-\mu)^r]$$

The first four moments about the mean:

$$\mu_1 = E(X-\mu) = 0$$
$$\mu_2 = E[(X-\mu)^2] = Var(X)$$
$$\mu_3 = E[(X-\mu)^3]$$
$$\mu_4 = E[(X-\mu)^4]$$

Moments about any point $\boldsymbol{a}$:

The r −th moment of a RV X about a point a is given by

$$\mu_r{}' = E[(X-a)^r]$$

The first four moments about a point $\boldsymbol{a}$:

$$\mu_1{}' = E(X-a) = E(X) - a = \mu - a$$
$$\mu_2{}' = E[(X-a)^2]$$
$$\mu_3{}' = E[(X-a)^3]$$
$$\mu_4{}' = E[(X-a)^4]$$

Moments about the origin (Raw moments):

The r −th moment of a RV X about the origin $(a = 0)$ is given by

$$\mu_r{}' = E(X^r)$$

The first four moments about the origin:

$$\mu_1{}' = E(X)$$
$$\mu_2{}' = E(X^2)$$
$$\mu_3{}' = E(X^3)$$
$$\mu_4{}' = E(X^4)$$

Relationship between raw and central moments

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

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

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

where

$$\mu_1' = E(X), \mu_2' = E(X^2), \mu_3' = E(X^3), \mu_4' = E(x^4)$$

Moments	Discrete	Continuous
$\mu_r = E[(X-\mu)^r]$	$\sum (x_i - \mu)^r p_i$	$\int_{-\infty}^{\infty} (x-\mu)^r f(x) dx$
$\mu_r{}' = E[(X-a)^r]$	$\sum (x_i - a)^r p_i$	$\int_{-\infty}^{\infty} (x-a)^r f(x) dx$
$\mu_r' = E(X^r)$	$\sum x_i^r p_i$	$\int_{-\infty}^{\infty} x^r f(x) dx$

Role of moments

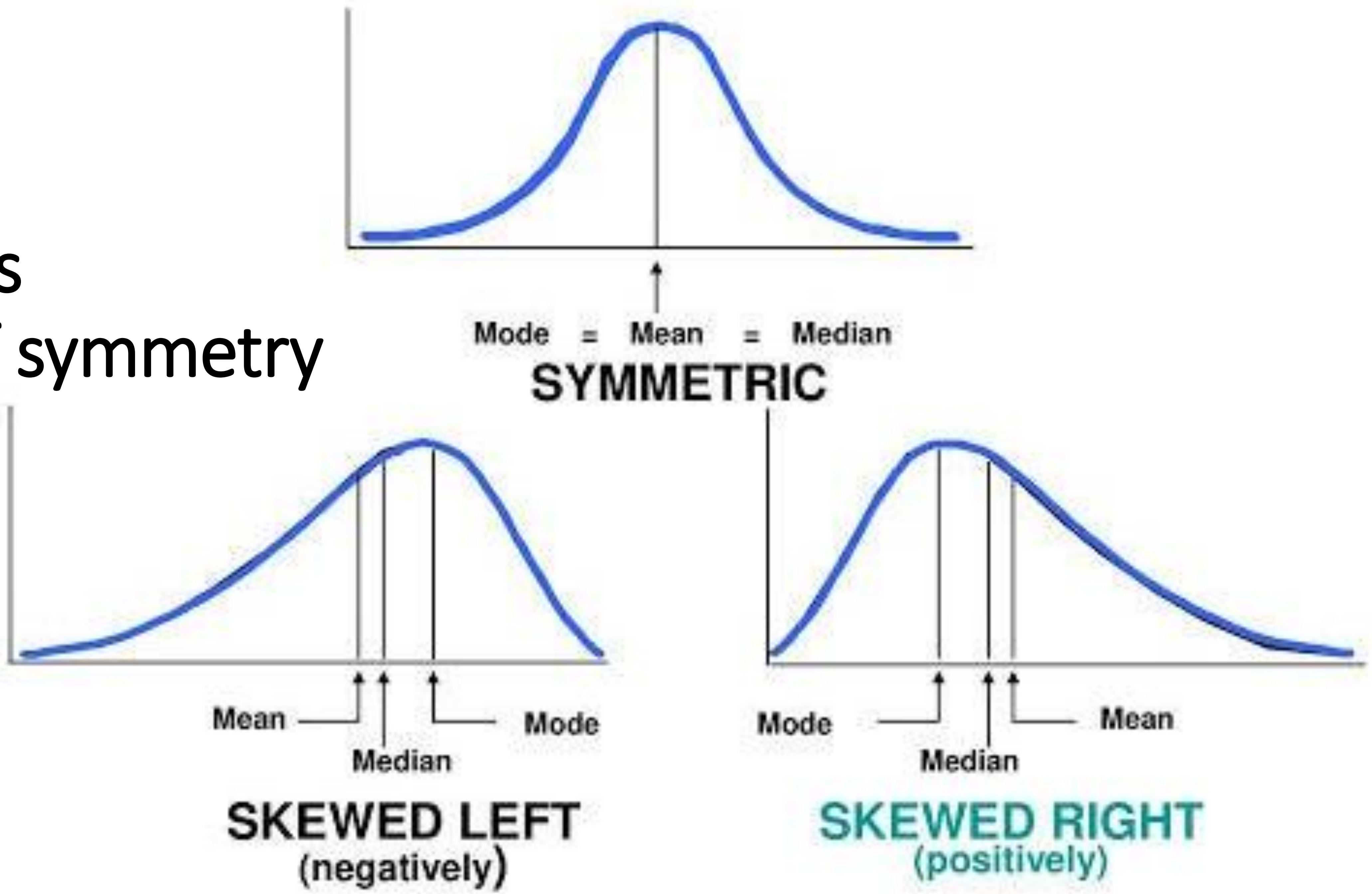
$\mu_1' -$ mean - centre of the probability function

$\mu_2 -$ variance - measures the dispersion of the distribution about the mean

$\mu_3 -$ skewness – measure the asymmetry

$\mu_4 -$ kurtosis – measure the degree of flatness (or) peakedness of the probability distribution near its centre.

Skewness
= Lack of symmetry



Coefficient of skewness

→ Find 'k'

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$$

→ $E(x)$

Raw $E(x^2)$

$E(x^3)$

$E(x^4)$

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

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

Central

→

→ Coefficients

Note:

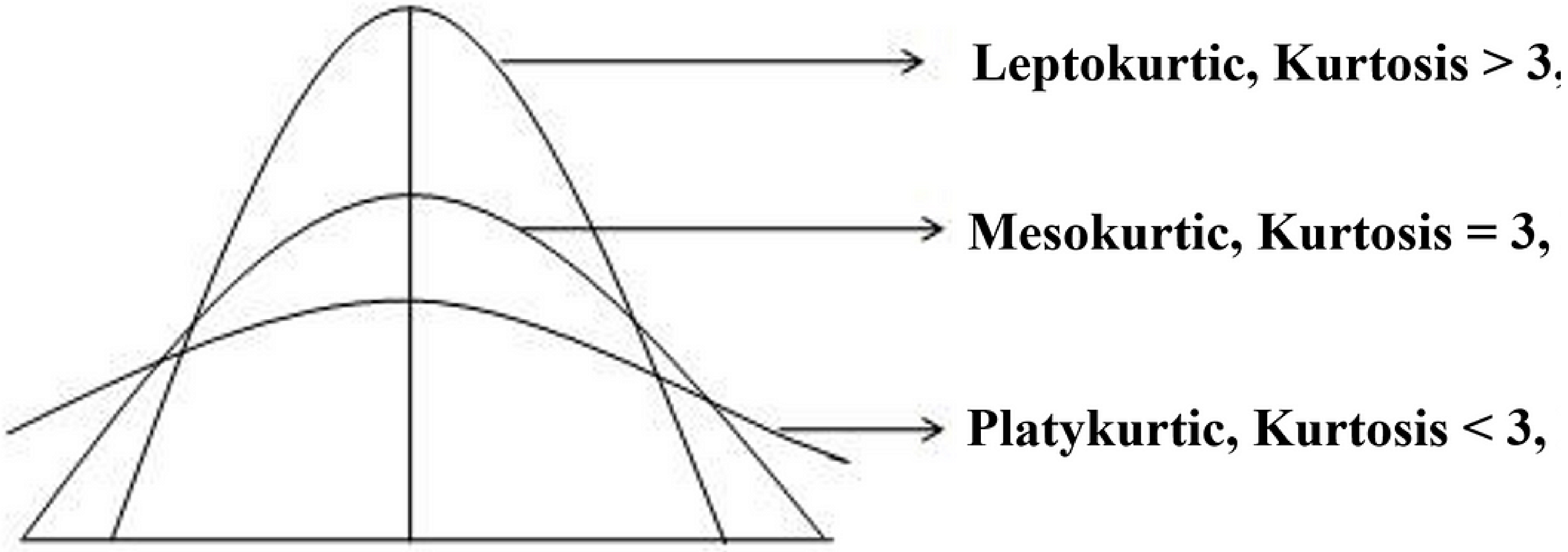
- For a symmetric distribution, odd moments about mean is 0.

$$\mu_1 = 0, \mu_3 = 0 \Rightarrow \gamma_1 = 0$$

- $\gamma_1 > 0 \Rightarrow$ positive skewness
- $\gamma_1 < 0 \Rightarrow$ negative skewness

Kurtosis = tailedness

- describes the shape of a probability distribution



Coefficient of kurtosis

$$\beta_2 = \frac{\mu_4}{\mu_2^2}$$

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

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

Types

(1) ***"Mesokurtic"***- If $\beta_2 = 3$, the curve is symmetric about the mean (normal; bell-shaped)

(2) ***"Platykurtic"*** - If $\beta_2 < 3$, the curve is more flat

(3) ***"Leptokurtic"*** - If $\beta_2 > 3$, the curve is more peaked

Example 8

A RV X has the following pmf

$X = x_i$	-2	-1	0	1	2	3
$P(X = x_i)$	0.1	k	0.2	$2k$	0.3	k

Find the value of the coefficient of skewness (γ_1) and the coefficient of kurtosis (β_2).

(Calculate these parameters first)

$$k = 0.1$$
$$E(X) = 0.8$$
$$E(X^2) = 2.8$$
$$E(X^3) = 4.4$$
$$E(X^4) = 14.8$$

The first moment about the origin

$$\mu_1' = E(X) = 0.8$$

The second moment about the origin

$$\mu_2' = E(X^2) = 2.8$$

The third moment about the origin

$$\mu_3' = E(X^3) = 4.4$$

The fourth moment about the origin

$$\mu_4' = E(X^4) = 14.8$$

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

$$= 2.8 - 0.8^2 = 2.16 \ \ (\mathrm{Var})$$

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

$$= 4.4 - 3(2.8)(0.8) + 2(0.8)^3$$

$$= -1.296$$

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

$$= 14.8 - 4(4.4)(0.8) + 6(2.8)(0.8)^2$$
$$- 3(0.8)^4$$
$$= 10.2432$$

Coefficient of skewness

$$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}} = -\frac{1.296}{(2.16)^{\frac{3}{2}}} = -0.4082$$

$\gamma_1 < 0 \Rightarrow$distribution is negatively skewed.

Coefficient of kurtosis

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{10.2432}{(2.16)^2} = 2.1956$$

$\beta_2 < 3 \Rightarrow$ distribution is platykurtic.

DIY

A RV X has the following pmf

$X = x_i$	1	2	3	4
$P(X = x_i)$	$k/3$	$k/6$	$k/3$	$k/6$

(i) Find the value of k

(ii) Compute mean, variance, β_1, β_2

Example 9

In a continuous distribution whose relative frequency density is given by

$$f(x) = y_0 x(2 - x), \qquad 0 \leq x \leq 2$$

where y_0 is a constant.

Find (i) Mean (ii) Variance (iii) β_1 (iv) β_2
(v) show that the distribution is symmetrical

Solution

To find y_0: $\int_{-\infty}^{\infty} f(x)dx = 1$

$$\int_0^2 y_0 x(2-x)dx = 1$$

$$y_0 \int_0^2 (2x - x^2)dx = 1$$

$$y_0 \left[2\frac{x^2}{2} - \frac{x^3}{3}\right]_0^2 = 1$$

$$y_0 \left[x^2 - \frac{x^3}{3}\right]_0^2 = 1$$

$$y_0\left[4-\frac{8}{3}\right]=1$$

$$y_0\left[\frac{4}{3}\right]=1 \Rightarrow y_0=\frac{3}{4}$$

r −th moment about origin:

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

$$= \int_0^2 x^r\, y_0 x(2-x)dx$$

$$= y_0 \int_0^2 x^{r+1}(2-x)dx$$

$$= \frac{3}{4}\int_0^2 (2x^{r+1} - x^{r+2})dx$$

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

$$= \frac{3}{4}\left[2\frac{(2)^{r+2}}{r+2} - \frac{2^{r+3}}{r+3}\right]$$

$$= \frac{3}{4}\left[\frac{2^{r+3}}{r+2} - \frac{2^{r+3}}{r+3}\right]$$

$$= \frac{3}{4} \times 2^{r+3}\left[\frac{1}{r+2} - \frac{1}{r+3}\right]$$

$$= 3 \times 2^{r+1}\left[\frac{1}{r+2} - \frac{1}{r+3}\right]$$

First four moments about origin:

$$r = 1 \Rightarrow \mu_1' = 3(2^2)\left[\frac{1}{3} - \frac{1}{4}\right] = 1$$

$$r = 2 \Rightarrow \mu_2' = 3(2^3)\left[\frac{1}{4} - \frac{1}{5}\right] = \frac{6}{5}$$

$$r = 3 \Rightarrow \mu_3' = 3(2^4)\left[\frac{1}{5} - \frac{1}{6}\right] = \frac{8}{5}$$

$$r = 4 \Rightarrow \mu_4' = 3(2^5)\left[\frac{1}{6} - \frac{1}{7}\right] = \frac{16}{7}$$

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

$$= \frac{6}{5} - 1^2 = \frac{1}{5}$$

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

$$= \frac{8}{5} - 3\left(\frac{6}{5}\right)(1) + 2(1)^3 = 0$$

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

$$= \frac{16}{7} - 4\left(\frac{8}{5}\right)(1) + 6\left(\frac{1}{5}\right)(1)^2 - 3(1)^4 = \frac{3}{35}$$

(i) Mean $\mu_1' = 1$

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

$$= \frac{6}{5} - 1^2 = \frac{1}{5}$$

(iii)

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{0}{1/5} = 0$$

(iv)

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3/35}{(1/5)^2} = \frac{15}{7} < 3$$

(v) To show that the distribution is symmetrical:

$$\mu_3 = 0 \Rightarrow \gamma_1 = 0$$

The coefficient of skewness is 0. So, the distribution is symmetrical.