One - dimensional random variable:

Ex: Consider the random experiment of tossing a coin. Here $S = \{H, T\}$.

Let us define a mapping $X : S \longrightarrow \mathbb{R}$ by

$$X(s) = \begin{cases} 1, & \text{if } s = H \\ 0, & \text{if } s = T \end{cases}$$

This real valued function X is called a random variable. Here range space = The set of all possible values of X = $R_x = \{0, 1\}$.

Def: Let E be an experiment & S is a sample space associated with the experiment. A function X assigning to every element $s \in S$, a real number $X(s)$, is called a random variable.

Discrete random variable:

Let X be a random variable. If the number of possible values of X [i.e R_X?] is finite or countably infinite, we call X a discrete r.v..

Suppose $x_1, x_2, x_3, \ldots$ are the possible values of X. with each possible outcome x_i we associate a number $p(x_i)$ or $P(X = x_i)$ called the probability of x_i. The numbers $p(x_i)$ $i = 1, 2, \ldots$ must satisfy the following conditions:

i) $p(x_i) \geq 0 \quad \forall i$

ii) $\sum_{i=1}^{\infty} p(x_i) = 1$

The function p defined above is called the probability function [or point probability function] of the random variable X.

The collection of pairs $(x_i, p(x_i))$ $i = 1, 2, \ldots$ is called the probability distribution of X.

Probability distribution table:

$X = x_i$	x_1	x_2	x_3	...
$P(X=x_i)$	$p(x_1)$	$p(x_2)$	$p(x_3)$	...

<u>Def:</u> Let X be a discrete r.v. with possible values $x_1, x_2, \ldots x_n, \ldots$. Then the expected value of X denoted by $E(X)$ or μ is defined as

$$E(X) = \sum_{i=1}^{\infty} x_i p(x_i)$$ if the series $\sum_{i=1}^{\infty} x_i p(x_i)$ converges absolutely i.e. $\sum_{i=1}^{\infty} |x_i p(x_i)| < \infty$. This number is also referred as the mean value of X.

Also $$E(X^2) = \sum_{i=1}^{\infty} x_i^2 p(x_i)$$

<u>Continuous random variable:</u>

~~Let~~ X is said to be a continuous r.v. if there exists a function f, called the probability density function (pdf) of X, satisfying the following conditions:

a) $f(x) \geq 0 \quad \forall x$

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

c) For any a, b with $-\infty < a < b < \infty$, we have

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

Expected value of X is given by

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

Also $E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\, dx$.

Def: Let X be a random variable. We define the variance of X denoted by $V(X)$ or σ_X^2 as

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

Def: The positive square root of $V(X)$ is called the standard deviation of X & is denoted by σ_X.

Cumulative Distribution Function:

Let X be a random variable. We define the cumulative distribution function (cdf) of the r.v. X as $F(x) = P(X \leq x)$.

a) If X is a discrete r.v. then

$F(x) = \sum_j p(x_j)$ where the sum is taken over all indices j satisfying $x_j \leq x$.

b) If X is a continuous r.v. then

$$F(x) = \int_{-\infty}^{x} f(s)\, ds.$$

Note:

1) The probability of X taking the value greater than t is defined by

$$P(X > t) = 1 - P(X \leq t)$$

2] If a & b are two specified values of X with $b > a$, then

$$P(a \le X \le b) = P(X \le b) - P(X \le a)$$

3] If $X = c$ where c is a constant then $E(X) = c$

4] If c is a constant & X is a random variable then $E(cX) = cE(X)$.

Problems:

1] If X is a random variable then prove that

$\cancel{V(X) = E(X^2) - [E(X)]^2}$

$$E(aX+b) = aE(X)+b, \quad \text{where } a, b \text{ are constants.}$$

~~Proof:~~

Proof: i) When X is a discrete r.v. then

$$
\begin{aligned}
E(aX+b) &= \sum_{i=1}^{\infty} (ax_i+b)\, p(x_i) \\
&= \sum_{i=1}^{\infty} ax_i\, p(x_i) + \sum_{i=1}^{\infty} b\, p(x_i) \\
&= a\sum_{i=1}^{\infty} x_i\, p(x_i) + b\sum p(x_i) \\
&= aE(X) + b
\end{aligned}
$$

ii) When X is a continuous r.v. then

$$
\begin{aligned}
E(aX+b) &= \int_{-\infty}^{\infty} (ax+b)\, f(x)\, dx \\
&= \int_{-\infty}^{\infty} ax\, f(x)\, dx + \int_{-\infty}^{\infty} b\, f(x)\, dx \\
&= a\int_{-\infty}^{\infty} x\, f(x)\, dx + b\int_{-\infty}^{\infty} f(x)\, dx \\
&= aE(X) + b
\end{aligned}
$$

2] If X is a random variable then prove that $V(aX+b) = a^2 V(X)$ where a, b are constants.

Proof W.K.T. $V(X) = E[X - E(X)]^2$

$$V(aX+b) = E[aX+b - E(aX+b)]^2$$
$$= E[aX + b - aE(X) - b]^2$$
$$= E[aX - aE(X)]^2$$
$$= E\left(a^2 [X - E(X)]^2\right)$$
$$= a^2 E[X - E(X)]^2$$
$$= a^2 V(X)$$

3] If X is a r.v. then prove that $V(X) = E(X^2) - [E(X)]^2$

Proof: $V(X) = E[X - E(X)]^2$ where $E(X) = \mu$

$$= E[X - \mu]^2$$
$$= E[X^2 + \mu^2 - 2\mu X]$$
$$= E(X^2) + E(\mu^2) - 2\mu E(X)$$
$$= E(X^2) + \mu^2 - 2\mu^2$$
$$= E(X^2) - \mu^2$$
$$= E(X^2) - [E(X)]^2$$

4] The probability distribution of a r.v. is given by the following table :

X	0	1	2	3	4	5	6
P(X)	K	3k	5k	7k	9k	11k	13k

i) For what value of k does this represent a valid probability distribution ?

ii) Find $P(X < 4)$

iii) Find $P(X \geq 5)$

iv) Find $P(3 < X \leq 6)$

v) Calculate mean, variance & std. deviation.

Soln: i) Given prob. distribution of a r.v. is valid when $\sum_{i=1}^{7} p(x_i) = 1$

$$p(x_1) + p(x_2) + p(x_3) + \cdots + p(x_7) = 1$$

$$k + 3k + 5k + 7k + 9k + 11k + 13k = 1$$

$$49k = 1$$

$$k = \frac{1}{49}$$

ii) $P(X < 4) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$

$$= \frac{1}{49} + \frac{3}{49} + \frac{5}{49} + \frac{7}{49}$$

$$= \frac{16}{49}$$

iii) $P(X \geq 5) = P(X=5) + P(X=6)$

$$= \frac{11}{49} + \frac{13}{49}$$

$$= \frac{24}{49}$$

iv) $P(3 < X \leq 6) = P(X=4) + P(X=5) + P(X=6)$

$$= \frac{9}{49} + \frac{11}{49} + \frac{13}{49}$$

$$= \frac{33}{49}$$

v) $E(X) = \sum_{i=1}^{7} x_i p(x_i)$

$$= 0\left(\frac{1}{49}\right) + (1)\left(\frac{3}{49}\right) + 2\left(\frac{5}{49}\right) + 3\left(\frac{7}{49}\right) + 4\left(\frac{9}{49}\right) + 5\left(\frac{11}{49}\right) + 6\left(\frac{13}{49}\right)$$

$$= \frac{203}{49} = 4.1428$$

$E(X^2) = \sum_{i=1}^{7} x_i^2 p(x_i)$

$$= 0^2\left(\frac{1}{49}\right) + 1^2\left(\frac{3}{49}\right) + 2^2\left(\frac{5}{49}\right) + 3^2\left(\frac{7}{49}\right) + 4^2\left(\frac{9}{49}\right) + 5^2\left(\frac{11}{49}\right) + 6^2\left(\frac{13}{49}\right)$$

$$= 19.857$$

$V(X) = E(X^2) - [E(X)]^2 = 2.694$

$\sigma_X = \sqrt{V(X)} = 1.641$

5] The prob. distribution of a r.v. is given by the following table:

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

i) Find k such that it represent a valid probability distribution.

ii) Find $E(X)$, $V(X)$, σ_X

iii) Find $P(X > -1)$

iv) Find $P(-1 \le X < 2)$

v) $F(2)$

Soln: i) Given prob. distribution is valid if

$$\sum_{i=1}^{6} p(x_i) = 1$$

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

$$0.6 + 4k = 1$$

$$\Rightarrow k = 0.1$$

ii) $$E(X) = \sum_{i=1}^{6} x_i p(x_i)$$

$$= (-2)(0.1) + (-1)(0.1) + 0 + (1)(0.2) + (2)(0.3) + 3(0.1)$$

$$= 0.8$$

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

$$E(x^2) = \sum_{i=1}^{6} x_i^2 p(x_i)$$

$$= (-2)^2(0.1) + (-1)^2(0.1) + 0 + (1)^2(0.2) + (2)^2(0.3) + (3)^2(0.1)$$

$$= 2.8$$

$$\therefore V(X) = 2.16$$

$$\sigma_X = 1.4696$$

iii) $$P(X > -1) = P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= 0.2 + 0.2 + 0.3 + 0.1$$

$$= 0.8$$

iv) $P(-1 \le X < 2) = P(X=-1) + P(X=0) + P(X=1)$

$= 0.1 + 0.2 + 0.2$

$= 0.5$

v) $F(x) = \sum_j p(x_j) \quad , \forall x_j \le x$

$F(2) = p(-2) + p(-1) + p(0) + p(1) + p(2)$

$= 0.1 + 0.1 + 0.2 + 0.2 + 0.3$

$= 0.9$

6] Suppose that a r.v. X takes values 1, 2, 3 ---- & $P(X=j) = \frac{1}{2^j}$, $j = 1, 2, 3$ ---- . Compute

i) P(X is even) ii) $P(X \geq 5)$ iii) P(X is divisible by 3)

Soln: i) $X = 2, 4, 6, ----$

$P(X \text{ is even}) = P(X=2) + P(X=4) + P(X=6) + ----$

$P(X \text{ is even}) = \frac{1}{2^2} + \frac{1}{2^4} + \frac{1}{2^6} + ----$

$S_\infty = \frac{a}{1-r}$

$P(X \text{ is even}) = \frac{1/2^2}{1-\frac{1}{2^2}} = \frac{\frac{1}{4}}{1-\frac{1}{4}} = \frac{\frac{1}{4}}{3/4} = \frac{1}{3}$

ii) $P(X \geq 5) = P(X=5) + P(X=6) + P(X=7) + ----$

$= \frac{1}{2^5} + \frac{1}{2^6} + \frac{1}{2^7} + ----$

$= \frac{\frac{1}{2^5}}{1-\frac{1}{2}} = \frac{\frac{1}{2^5}}{1/2} = \frac{1}{2^5} \times 2 = \frac{1}{2^4} = \frac{1}{16}$

iii) P(X is divisible by 3) $= P(X=3)+P(X=6)+P(X=9)+\cdots$

$$= \frac{1}{2^3}+\frac{1}{2^6}+\frac{1}{2^9}+\cdots$$

$$= \frac{\frac{1}{2^3}}{1-\frac{1}{2^3}} = \frac{\frac{1}{8}}{8-1/8} = \frac{1}{8}\times\frac{8}{7}=\frac{1}{7}$$

7) A box contains 12 items of which 4 are defective. A sample of 3 items is selected from the box. Let X denote the no. of defective items in the sample. Find the probability distribution of X. Determine the mean & std. deviation.

Soln: Out of 3 items selected, none can be defective, one can be defective, two can be defective, three can be defective.

$\therefore$ X can take values 0, 1, 2, 3

X	0	1	2	3
P(X)	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$

$$P(X=0) = P(\text{no defective}) = \frac{{}^8C_3}{{}^{12}C_3} = \frac{56}{220}$$

$$P(X=1) = P(\text{one defective \& two non defective}) = \frac{{}^4C_1\times{}^8C_2}{{}^{12}C_3}$$

$$= \frac{112}{220}$$

$P(X=2) = P(2 \text{ defective } \& \text{ 1 non defective})$

$$= \frac{{}^4C_2 \times {}^8C_1}{{}^{12}C_3} = \frac{48}{220}$$

$P(X=3) = P(\text{all are defective})$

$$= \frac{{}^4C_3}{{}^{12}C_3} = \frac{4}{220}$$

$$E(X) = \sum_{i=1}^{4} x_i p(x_i)$$

$$= 0 + (1)\left(\frac{112}{220}\right) + (2)\left(\frac{48}{220}\right) + (3)\left(\frac{4}{220}\right)$$

$$= 1$$

$$E(X^2) = \sum_{i=1}^{4} x_i^2 p(x_i)$$

$$= 0 + (1)^2\left(\frac{112}{220}\right) + 2^2\left(\frac{48}{220}\right) + 3^2\left(\frac{4}{220}\right)$$

$$= \frac{340}{220} = 1.5454$$

$$V(X) = E(X^2) - [E(X)]^2 = 0.5454$$

$$\sigma_X = \underline{0.7385}$$