One Dimensional Random Variables

Definition:

A random variable is a function that assigns a real number

$X(s) = x$ to every element $s \epsilon S$, where S is the sample space corresponding to a random experiment E.

Range Space R_X: is the set of possible values of X, is a subset of real numbers R.

Random variables are denoted by capital letters X, Y, Z etc and x, y, z etc denotes one of its values.

Example: Suppose that we toss two coins and consider the sample space associated with this experiment.

$$S = \{HH, HT, TH, TT\}$$

Define a random variable X as follows:

X is the number of heads obtained in the two tosses.

Hence $X(HH) = 2,\ X(HT) = 1,\ X(TH) = 1, X(TT) = 0$

$$R_X = \{0, 1, 2\}$$

Equivalent event: Let X be a random variable defined on S and let R_X be its range space. Let B be an event with respect to R_X. (i.e., $B \subset R_X$). Suppose that $A = \{s \in S$ / X(s)∈ $B\}$ [i.e., A consists of all outcomes in S for which X(s)∈ B.]. Then we say that A and B are equivalent.

Example: Consider the tossing of two coins

$$S = \{HH, HT, TH, TT\}$$

X is the number of heads obtained in the two tosses and $R_X = \{0, 1, 2\}$

Let B={1}, and A={HT, TH}

$$P(B) = P(A) = 2/4 = 0.5$$

Similarly, B={1, 2} then A= $\{HH, HT, TH\}$ and Hence $P(B) = P(A) = \frac{3}{4}$

Types of Random Variables:

1. Discrete Random variable
2. Continuous Random variable

1. Discrete Random variable: If X is a Random variable which can take finite number or countably infinite number of values, then X is called a Discrete Random variable.

Probability function or Probability mass function (pmf):

If X is a Discrete Random variable which can take the values $x_1, x_2, x_3, \ldots \, such \; that \; P(X = x_i) = p_i$, then p_i is called Probability function or Probability mass function (pmf), provided p_i (i=1, 2, 3, …) satisfy the following conditions.

i). $p_i \geq 0 \;\; for \; all \; i \quad (p(x_i) \geq 0 \; for \; all \; i)$

ii). $\sum_i p_i = 1 \quad (\sum_{i=1}^{\infty} p(x_i) = 1)$

2. Continuous Random variable: If X is a random variable which can take all the values (i.e., infinite number of values) in an interval, then X is called Continuous Random variable.

Probability Density Function (pdf):

If X is continuous random if there exists a function $f(x)$, called the Probability Density Function (pmf) provided $f(x)$ satisfy the following conditions.

i). $f(x) \geq 0$ for all $x \in R_X$

(ii). $\int_{-\infty}^{\infty} f(x) = 1$

(iii). For any a, b with $-\infty < a < b < \infty$,

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

Cumulative Distribution function (cdf):

If X is a random variable discrete or continuous, then $P(X \leq x)$is called the Cumulative Distribution function (cdf) of X or Distribution function of X and denoted as $F(x)$.

If X is Discrete : $F(x) = \sum_{x_j \leq x} p(x_j)$

If X is Continuous : $F(x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(x)dx$

Properties of cdf $F(x)$:

1. $F(x)$ is non-decreasing function of x

i.e., if $x_1 < x_2, then\ F(x_1) < F(x_2)$

2. $F(-\infty) = 0\ and\ F(\infty) = 1$

3. $P(a < X < b) = \int_a^b f(x)dx = \int_{-\infty}^b f(x)dx - \int_{-\infty}^a f(x)dx$

$$= F(b) - F(a)$$

4. $f(x) = \frac{d}{dx}[F(x)] = F'(x)$ at all the points where $F(X)$ is differentiable.

Mean: Let X be **a discrete random** variable.

Then mean value of X denoted by $E(X)$, defined as,

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

- If $g(X)$ is a function of X, then expectation of $g(X)$ is,

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

Let X be **a continuous random** variable with pdf $f(x)$.

Then expected value of X is defined as,

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

- If $h(X)$ is a function of X, then expectation of $h(X)$ is,

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

Properties:

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

$E\left(E(X)\right) = E(X)$

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

Variance: Let X be a random variable. We define the **variance of X**, denoted by $V(X)$ or σ^2 as follows,

$$V(X) = \sigma^2 = E[X - E(X)]^2 = E\left[X^2 + \left(E(X)\right)^2 - 2XE(X)\right]$$

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

The positive square root of $V(X)$is called the standard deviation of X and is denoted by σ or σ_X

Properties:

$V(X + c) = V(X)$, where c is a constant

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

Markov's Inequality:

If X is a random variable assuming non negative values then $P(X \geq a) \leq \frac{E(X)}{a}$,

$a \neq 0, a > 0$

Chebyshe's Inequality (measure for bound):

Let X is a random variable with mean $E(X) = \mu$ and standard deviation σ, then for any $k > 0$,

$$P(\mu - k\sigma < X < \mu + k\sigma) \geq 1 - \frac{1}{k^2}$$

i.e.,

$$P(|X - \mu| < k\sigma) \geq 1 - \frac{1}{k^2}$$

$$P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}, \qquad k > 0$$

$$P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{1}{k^2}$$

Using Markov’s inequality with X as $(X-\mu)^2$ and a as $k^2\sigma^2$

$$P[(X-\mu)^2 \geq k^2\sigma^2] \leq \frac{E[(X-\mu)^2]}{k^2\sigma^2} = \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$$

Therefore, Chebyshe’s inequality,

$$P(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}$$

1) Use the Chebyshe's inequlity to find the following

$(a) P(|X - \mu| < 3)$

$(b) P(-2 < X < 8)$

$(c) P(|X - \mu| \geq 3)$

Where X is a random variable, $E(X) = 3,\ E(X^2) = 13$

2) Suppose that it is known that the number of items produced by a factory during a week is a random variable with mean value 50.

a) What can be said about the probability that this week’s production, which exceed 75?

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

3) The number of patients requiring ICU in a hospital is a random variable with mean 18 and standard deviation 2.5. Determine the probability that the number of patients between 8 and 28.