Introduction to Probability and Statistics
Course ID:MA2203

Lecture-4

Course Teacher: Dr. Manas Ranjan Tripathy

Department of Mathematics
National Institute of Techonology, Rourkela

- Random Variable: A random variable X is finite and single valued function from the sample space S to $\mathcal{R}$, such that the inverse images under X of all Borel sets in $\mathcal{R}$ are events. That is $X^{-1}(B) = \{w : X(w) \in B\}$ event for all $B \in \mathcal{B}$. Class of Borel sets is the collection of open or closed intervals in $\mathcal{R}$, which is closed under countable union, countable intersection and complementation. In order to verify that a real valued function on S is a random variable, it is not necessary to check for all Borel sets. It is sufficient to verify the condition for any class of subsets of $\mathcal{R}$. Here we take the class of semi-closed intervals $(-\infty, x]$ $x \in \mathcal{R}$. Note that, for any real a the probability $P(X = a)$ with which X assumes a is defined. and for any interval I, the probability $P(X \in I)$ is defined.
- We can see that, the semi-closed interval

$$(-\infty, x] = \bigcap_{n=1}^{\infty} (-\infty, x + \frac{1}{n}).$$

- Examples (i): Suppose we toss a coin once. Here $S = \{H, T\}$. Let us define a function $X :\rightarrow \mathcal{R}$, such that X is the number of heads turns up. To verify that it is a random variable, observe that $X(T) = 0$, $X(H) = 1$. Take a subset of $\mathcal{R}$ as $(-\infty, x]$, $x \in \mathcal{R}$.

$$\begin{aligned} X^{-1}(-\infty, x] &= \emptyset, \ \textit{if } x < 0, \\ &= \{T\}, \ \textit{if } 0 \leq x < 1, \\ &= \{T, H\}, \ \textit{if } x \geq 1. \end{aligned}$$

In all the cases $X^{-1}(-\infty, x]$ is an event. Hence X is a random variable.

- (ii): Suppose we throw a die once. $S = \{1, 2, 3, 4, 5, 6\}$. Define X as the number shows up when we throw. That is $X(1) = 1$, $X(2) = 2$, $X(3) = 3$, $X(4) = 4$, $X(5) = 5$, $X(6) = 6$. To check whether X is a random variable, observe that,

$$
\begin{aligned}
X^{-1}(-\infty, x] &= \emptyset,\ \textit{if } x < 1, \\
&= \{1\},\ \textit{if } 1 \leq x < 2, \\
&= \{1, 2\},\ \textit{if } 2 \leq x < 3 \\
&= \{1, 2, 3\}\ \textit{if } 3 \leq x < 4 \\
&= \{1, 2, 3, 4\}\ \textit{if } 4 \leq x < 5 \\
&= \{1, 2, 3, 4, 5\}\ \textit{if } 5 \leq x < 6 \\
&= \{1, 2, 3, 4, 5, 6\},\ \textit{if } x \geq 6.
\end{aligned}
$$

  In all the cases $X^{-1}(-\infty, x]$ is an event. Hence X is a random variable.

- In general, we can say that the random variable is the quantity that we observe in a random experiment. The number of heads, the number shows up in throwing a die, the number of deaths by cancer, the number of accidents in a city, amount of rain fall, hardness of steel, etc.

- Distribution Function or Cumulative Distribution Function (CDF): A function $F(x)$ which is defined in $(-\infty, \infty)$ such that it is monotonically non-decreasing, right continuous and $F(-\infty) = 0$, $F(\infty) = 1$. The CDF of a random variable X is defined as $F(x) = P(X \leq x)$, we read it as the probability that the random variable X will not exceed x. Here $x \in \mathcal{R}$.
  (i) The probability that the random variable X will be in the interval $a < X \leq b$ is computed as $P(a < X \leq b) = F(b) - F(a)$. The interval $(-\infty, b]$ is the disjoint union of $(-\infty, a]$ and $(a, b]$. Hence $F(b) = P(X \leq a) + P(a < X \leq b)$.
- Types of random variables: (i) Discrete type, (ii) Continuous type.
- Discrete Type RV: A random variable X is said to be discrete if X assumes only finitely or countable number of values, say $x_1, x_2, \ldots,$ called the possible values of X with probabilities $p_1 = P(X = x_1)$, $p_2 = P(X = x_2), \ldots$ whereas $P(X \in I) = 0$ for any interval I that does not contain any x_i. Here $p_i > 0$ and $\sum_i p_i = \sum_i P(X = x_i) = 1$ and these p_is are known as the probability mass function (pmf) of X. The CDF of a discrete type random variable X is obtained as

$$F(x) = \sum_{x_i \leq x} P(X = x_i).$$

  Moreover $P(a < X \leq b) = \sum_{a<x\leq b} P(X = x)$, $P(a < X < b) = \sum_{a<x<b} P(X = x)$.

- Examples of Discrete Type RV: (i) If we toss a coin once, then $S = \{H, T\}$. Let X be the number of tails. Then $X(H) = 0 = x_1$, $X(T) = 1 = x_2$. Further $p_1 = P(X(H) = 0)$, $p_2 = P(X(T) = 1)$ is the probability mass function of X. Here we have two points. Also we have $p_1 + p_2 = 1$. If the coin is fair we can take $p_1 = p_2 = 1/2$. (ii) Tossing of a die, X is the number that shows up. (iii) Suppose we toss two coins simultaneously, X is the sum of head and tail. (iv) Suppose we throw two fair die simultaneously, X be the sum of two numbers that shows up.
- Continuous Type RV: A random variable X and its distribution are called continuous if if its distribution function $F(x)$ can be obtained by an integral,
$$F(x) = \int_{-\infty}^{x} f(v)dv,$$
here $f(x) > 0$ and is known as the probability density function (pdf) of X, and
$$\int_{-\infty}^{\infty} f(x)dx = 1.$$
Now differentiating this $F(x)$ at the point of continuity, we have
$$F'(x) = f(x).$$
- Moreover, we have $P(a < X < b) = P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = \int_a^b f(x)dx$.

- Examples of Continuous Type RV: (i) Let X have the density function $f(x) = 0.75(1 - x^2)$, if $-1 \leq x \leq 1$ and zero otherwise. Find the distribution function. Find the probabilities $P(-\frac{1}{2} \leq X \leq \frac{1}{2})$, $P(\frac{1}{4} \leq X \leq 2)$.
  Ans: To obtain the CDF, $F(x)$ we have

$$\begin{aligned} F(x) &= 0,\ \textit{if } x \leq -1, \\ &= \int_{-\infty}^{x} 0.75(1 - v^2)dv \\ &\quad = 0.5 + 0.75x - 0.25x^3,\ \textit{if } -1 < x \leq 1 \\ &= 1,\ \textit{if } x > 1. \end{aligned}$$

  Now $P(-\frac{1}{2} \leq X \leq \frac{1}{2}) = \int_{-\frac{1}{2}}^{\frac{1}{2}} f(x)dx = 0.6875$. $P(\frac{1}{4} \leq X \leq 2) = \int_{\frac{1}{4}}^{2} f(x)dx = 0.3164$.

- Some More Examples of Continuous Type RV: (i) The probability density function of a random variable X is

$$f(x) = \begin{cases} \frac{sinx}{2}, & \text{if } 0 \leq x \leq \pi \\ 0, & \text{elsewhere.} \end{cases}$$

  Check that it is a probability density function and find its cumulative distribution function. Further find $P(1/2 < X < \pi)$ and $P(X > \pi/2)$.

- (ii) Let X be a random variable having probability density function

$$f(x) = \begin{cases} \frac{x}{6} + k, & \text{if } 0 \leq x \leq 3 \\ 0, & \text{elsewhere.} \end{cases}$$

  Finf the value of k, and obtain the cumulative distribution function $F(x)$. Further find (a) $P(1 < X < 2)$ (b) $P(X > 1.8)$ (c) $P(3/2 < X < 3)$.

- (iii) A continuous random variable have the probability density function

$$f(x) = \begin{cases} ke^{-kx}, & \text{if } x > 0 \\ 0, & \text{elsewhere.} \end{cases}$$

  Find the value of k and the cumulative distribution function. Further obtain the probabilities (a) $P(X > 1/2)$ (b) $P(1 < X < 2)$ and (c) $P(X < 10)$.

- Some More Examples of Discrete Type RV: (i) Let X be a discrete type random variable having probability mass function given by

x	0	1	2	3	4	5	6	7
$P(X = x)$	0	k	$2k$	$2k$	$3k$	k^2	$2k^2$	$7k^2 + k$

  Find the value of k and the cumulative distribution function of X. Further find (a) $P(0 < X < 1.5)$ (b) $P(X \geq 5)$ (c) $P(1.9 \leq X < 6)$ (d) $P(X < 8)$

- (ii) Suppose we toss pair of dice simultaneously. Let X denotes the minimum of two numbers that appear. Show that X is a random variable and find its cumulative distribution function $F(x)$. Do the same problem if X denotes the maximum of two numbers.

- (iii) Let X be the sum of two numbers that appear when two dice are thrown simultaneously. Show that X is random variable and also obtain its cumulative distribution function.

- (iv) Suppose we toss 3 coins simultaneously. Let X denotes the sum of the number of heads. Find the cumulative distribution function if X is a random variable.