3.1 – 3.2

Random variables, Probability distribution of discrete random variable

The concept of random variable allows us to pass from the experimental outcomes themselves to a numerical function of the outcomes.

A probability distribution, shows the probabilities of events in an experiment. The quantity that we observe in an experiment will be denoted by X and called a random variable or stochatic variable.

Random variables (Definition)

For a given sample space S of some experiment, a random variable (rv) is any rule that associates a number with each outcome in S. In mathematical language, a random variable is a function whose domain is the sample space and whose range is the set of all real numbers.

i.e. $X : S \to R$

$X(s) = x$ means that x is the value associated with the outcome s by the r.v X.

Example:

Consider an experiment in which 9-volt batteries are tested until one with an acceptable voltage(S) is obtained. The Sample space is $S = \{S, FS, FFS, FFFS, \ldots\}$

Define random variable X = no. of batteries before the experiment terminates

then $X(S) = 1$, $X(FS) = 2$,
$X(FFS) = 3$ ---, $X(FFFFFFS) = 7$, and so on. (Infinite)

Scanned by CamScanner

Two types of random variables:

(a) Discrete Random Variable:

A discrete random variable is an r.v. whose possible values either constitute a finite set or else can be listed in an infinite sequence in which there is a 1st element, a 2nd element, and so on ("countably" infinite)

Examples If we count i) Cars on a road.
ii) defective screws in a production

(b) Continuous Random variable:

A random variable is continuous if both of the following apply:

1. Its set of possible values consists either of all numbers in a single interval on the number line (e.g., $-\infty$ to ∞) or all numbers in a disjoint union of such intervals (e.g., $[0,10] \cup [20,30]$).
2. No possible value of the variable has positive probability, that is $P(X = c) = 0$ for any possible value c.

Examples :-

If we measure i) electric voltage
ii) rainfall

The probability distribution/probability mass function (pmf):

The probability mass function (pmf) of a discrete r.v. is defined for every x by

$$p(x) = P(X=x) = P\{\text{all } s \in S : X(s) = x\}$$

For any pmf the following two conditions are required

i) $p(x) \geq 0$,

ii) $\sum_{x} p(x) = 1$.

Cumulative distribution function (cdf):

The cumulative distribution function (cdf) $F(x)$ of a discrete r.v. variable X with pmf $p(x)$ is defined by

$$F(x) = P(X \leq x) = \sum_{y : y \leq x} p(y)$$

For any number x, $F(x)$ is the probability that the observed value of X will be at most x.

Proposition:

$$P(a < X \leq b) = F(b) - F(a) = \sum_{a < x \leq b} p(x)$$, X = discrete r.v.

Scanned by CamScanner

Example:

Let us consider the random experiment of tossing two coins

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

Let the random variable X = " the number of heads "

Then $X : S \to R$ defined by

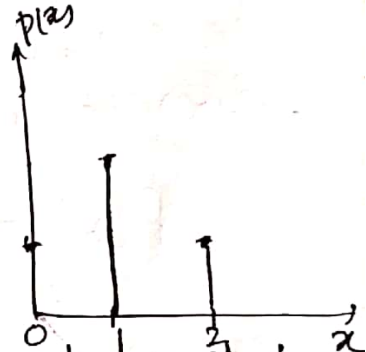
$X(HH) = 2$
$X(HT) = 1$
$X(TH) = 1$
$X(TT) = 0$

Thus the spectrum of X is $\{0, 1, 2\}$

Here the event $(X = 1)$ = " one head "
$= \{HT, TH\}$

Also $P(X=0) = \frac{1}{4}$, $P(X=1) = \frac{2}{4} = \frac{1}{2}$

$P(X=2) = \frac{1}{4}$



line graph of pmf

The pmf is given by

x:	0	1	2
$p(x)$:	$1/4$	$1/2$	$1/4$

clearly $\sum p(x) = 1$

Now, if $x < 0$, $F(x) = P(X \le x) = 0$

$0 \le x < 1$, $F(x) = P(X \le x) = \frac{1}{4}$

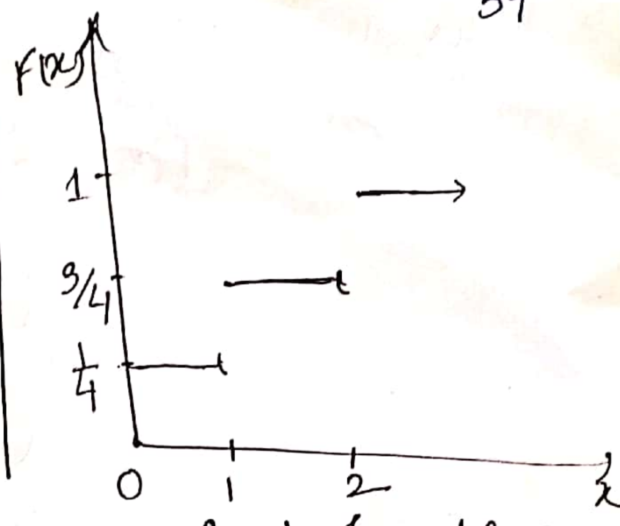
$1 \le x < 2$, $F(x) = P(X \le x) = P(X < 2)$
$= P(X=0) + P(X=1)$
$= \frac{1}{4} + \frac{1}{2} = \frac{1+2}{4} = 3/4$

$2 \le x < \infty$, $F(x) = P(X < \infty)$
$= P(X=0) + P(X=1) + P(X=2)$
$= 1/4 + 1/2 + 1/4 = 1$

Scanned by CamScanner

Thus the cdf is given by

$$F(x) = \begin{cases} 0, & -\infty < x < 0 \\ \frac{1}{4}, & 0 \le x < 1 \\ 3/4, & 1 \le x < 2 \\ 1, & 2 \le x < \infty \end{cases}$$



Graph of cdf F(x)

Example:

A store carries flash drives with either 1 GB, 2 GB, 4 GB, 8 GB or 16 GB of memory. The accompanying table gives the distribution of X = the amount of memory in a purchased drive:

x :	1	2	4	8	16
$p(x)$:	0.05	0.10	0.35	0.40	0.10

Let first determine F(x) for each of the 5 possible values of X:

$F(1) = P(X \le 1) = P(X = 1) = p(1) = 0.05$

$F(2) = P(X \le 2) = P(X = 1 \text{ or } 2) = p(1) + p(2) = 0.05 + 0.10 = 0.15$

$F(4) = P(X \le 4) = p(1) + p(2) + p(4) = 0.05 + 0.10 + 0.35 = 0.50$

$F(8) = P(X \le 8) = p(1) + p(2) + p(4) + p(8) = 0.90$

$F(16) = P(X \le 16) = 1$

Thus the cdf is

$$F(x) = \begin{cases} 0, & x < 1 \\ 0.05 & 1 \le x < 2 \\ 0.15 & 2 \le x < 4 \\ 0.50 & 4 \le x < 8 \\ 0.90 & 8 \le x < 16 \\ 1 & 16 \le x \end{cases}$$