

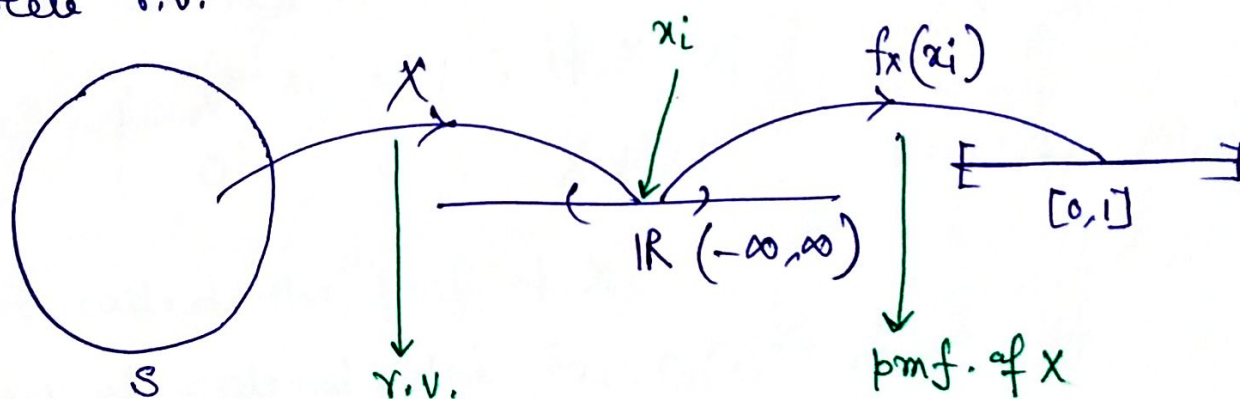
Random variable, Distribution function, Probability mass function

Random variable :

Let S be a sample space associated with a given random experiment E .

A real valued (measurable) function defined on S is called a one-dimensional random variable (r.v.).

If the sample space is discrete, then the corresponding random variable is said to be discrete r.v.



Example:- Let us consider the example of tossing a unbiased coin

$$S = \{H, T\}$$

Let us define the real valued fn. X on S as

$X = 1$, H appears

$= 0$, T appears.

$X \rightarrow$ discrete r.v.

Probability mass function

Let X be a one-dimensional discrete random variable taking at most countably infinite number of values x_1, x_2, \dots (i.e., $x_i, i \in I$, I being the index set) then its probabilistic behaviour at each real point is described by a function, called the probability mass function (pmf) and is defined as follows.

Definition: If X is a discrete r.v with distinct values $x_1, x_2, \dots, x_n, \dots$ then the function $p(x)$ or $p_x(x)$ or $f_x(x)$ is defined as follows

$$p_x(x) = \begin{cases} P(X=x_i) = p_i & ; \text{ if } X=x_i \\ 0 & ; X \neq x_i \end{cases} \quad ; i \in I. \\ I = \{1, 2, \dots, n, \dots\}$$

and is called the pmf of X .

The set of ordered pairs $\{x_i, p_x(x_i)\}, i \in I$, or, $\{(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n), \dots\}$ represents

the probability distribution of the r.v X .

If $p_x(x_i); i \in I$ is the pmf, then it must satisfy the following two conditions.

(i) $p_x(x_i) \geq 0 \quad \forall i \in I.$

(ii) $\sum_{i \in I} p_x(x_i) = 1.$

Example: let us consider the example taken in Lecture 4 P(1)

$$\left. \begin{aligned} P[X=1] &= \frac{1}{2} = p_1 \\ P[X=0] &= \frac{1}{2} = p_2 \end{aligned} \right\} \begin{array}{l} \text{pmf} \\ \text{of } X. \end{array}$$

Cumulative distribution fn. (cdf) or distribution fn. (df) of a discrete r.v. ;

Let $P: \Delta \rightarrow \mathbb{R}$ be a probability function, where Δ is the class of subsets of S (S being the sample space associated with the r.v. X) or power set of S forming the class of events. Then we remember that the order 3 tuple (S, Δ, P) is called the probability space. Let X be the random variable defined on the event space S .

The distribution fn. or cumulative distribution fn. of the r.v. X w.r.t the prob. space (S, Δ, P) is a real valued fn. $F_X(x)$ of real variable x , defined in $(-\infty, \infty)$ such that

$$F_X(x) = P(-\infty < X \leq x) \quad \forall x \in (-\infty, \infty).$$

It is clear that the range of the fn. is a subset of $[0, 1]$.

Let X be a discrete r.v.,

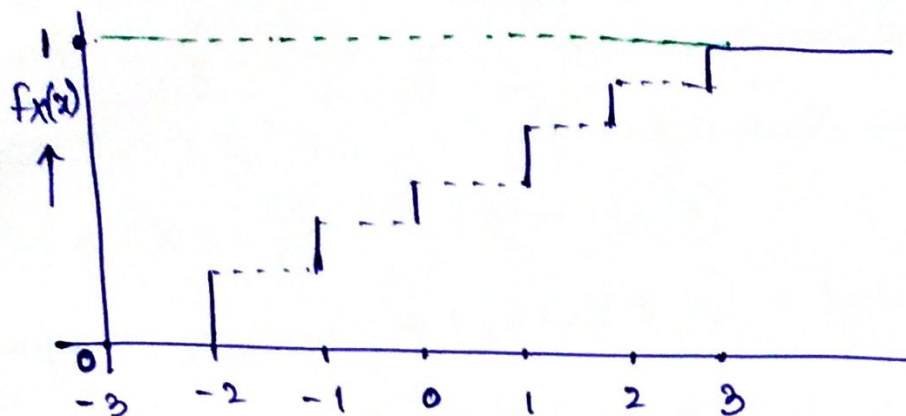
then cdf of X is defined as

$$F_X(x) = \sum_{x_j \leq x} P(X = x_j) = \sum_{j=-\infty}^i p_j \quad \text{if } x_i \leq x < x_{i+1},$$

$i = 0, \pm 1, \pm 2, \pm 3, \dots$

Thus $F_X(x)$ is a step fn.

See the fig. in next page.



$X \rightarrow$ discrete r.v.
 \hookrightarrow p.m.f \downarrow
 $P(X=x_i) = p_i, i = -2, -1, 0, 1, 2, 3$

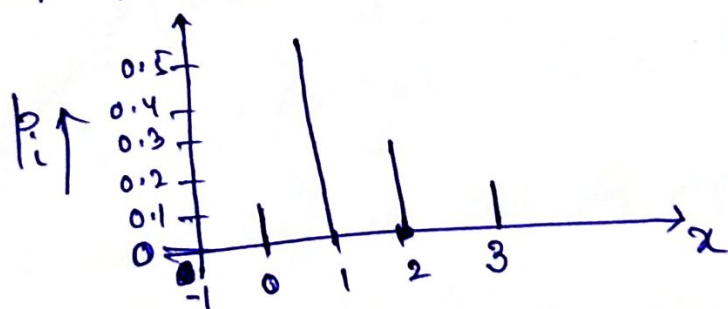
Fig. cdf of a discrete distribution

Example: Let X be a discrete r.v. s.t

$$P(X=0)=0.1, P(X=1)=0.5, P(X=2)=0.25, P(X=3)=0.15$$

Hence, $\sum_{i=0}^3 P(X=i) = 1.$

If pmf look like as given in the following fig



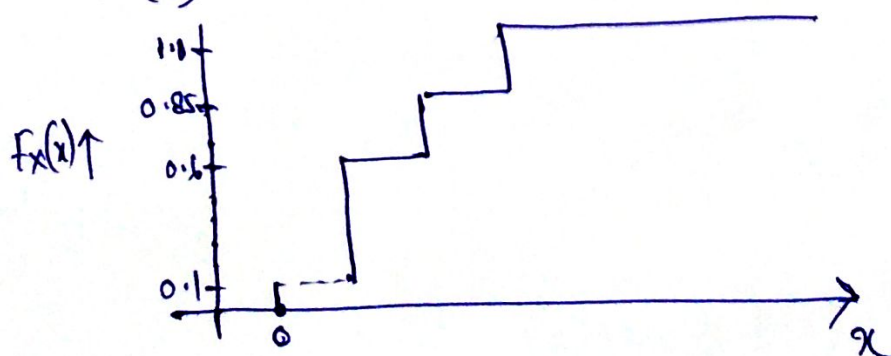
Let us ~~calculate~~ calculate its cdf.

$$F_X(0) = P(X \leq 0) = P(X=0) = 0.1$$

$$F_X(1) = P(X \leq 1) = P(X=0) + P(X=1) = 0.1 + 0.5 = 0.6$$

$$F_X(2) = P(X \leq 2) = P(X=0) + P(X=1) + P(X=2) = 0.1 + 0.5 + 0.25 = 0.85$$

$$F_X(3) = P(X \leq 3) = p_0 + p_1 + p_2 + p_3 = 1.$$



Properties of distribution fn. or cumulative distribution fn.

1. Let $F_X(x)$ be the cdf of the r.v. X and $a < b$, then

$$P(a < X \leq b) = F_X(b) - F_X(a)$$



Proof: $P(X \leq a) + P(a < X \leq b) = F_X(b)$,
by defn.

$$\Rightarrow P(a < X \leq b) = F_X(b) - P(X \leq a) \\ = F_X(b) - F_X(a).$$

$$\text{Cor. 1. } P(a \leq X \leq b) = P[(X=a) \cup (a < X \leq b)] \\ = P(X=a) + P(a < X \leq b) \quad \because \text{disjoint.} \\ = P(X=a) + F_X(b) - F_X(a).$$

$$\text{Cor. 2. } P(a < X < b) \neq P(X=b) = P(a < X \leq b) \\ \Rightarrow P(a < X < b) = F_X(b) - F_X(a) - P(X=b).$$

2. If $F_X(x)$ be the cdf of one-dimensional r.v. X ,

then (i) $0 \leq F_X(x) \leq 1$

(ii) $F_X(x) \leq F_X(y)$ if $x < y$.

$$3. \quad F(-\infty) = \lim_{x \rightarrow -\infty} F_X(x) = 0$$

$$F(\infty) = \lim_{x \rightarrow \infty} F_X(x) = 1.$$