

* Probability distributions

Random variable:- A set of possible values from a random experiment.

→ A random variable, value is unknown.

→ A function that assigns values to each of experiment outcomes.

tossing a coin - $\{H, T\}$

$$X = \{0, 1\}$$

$$\begin{cases} P(X=H) = \frac{1}{2} \\ P(X=T) = \frac{1}{2} \end{cases}$$

→ A function

$$= \frac{1}{n} \quad (\text{where } n \text{ is total no of outcomes})$$

$$P(H) = \frac{1}{2}$$

* dice - 1, 2, 3, 4, 5, 6.

$$\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6} \dots \frac{1}{6}$$

$$\Rightarrow \frac{1}{n} \quad \text{where } \underline{n=6}$$

↓
function that can be used to get probability.

Outcomes of Experiment

→ tossing a coin
→ Throwing a dice
(discrete outcomes)

↓
Probability mass function

→ calculate the prob if a student height is below 170 cm.

$$160, 160.1, 160.2$$

(continuous outcomes)

↓
Probability density function

* irrespective of outcome nature, draw outcomes in a form of distribution \Rightarrow probability distribution functions.

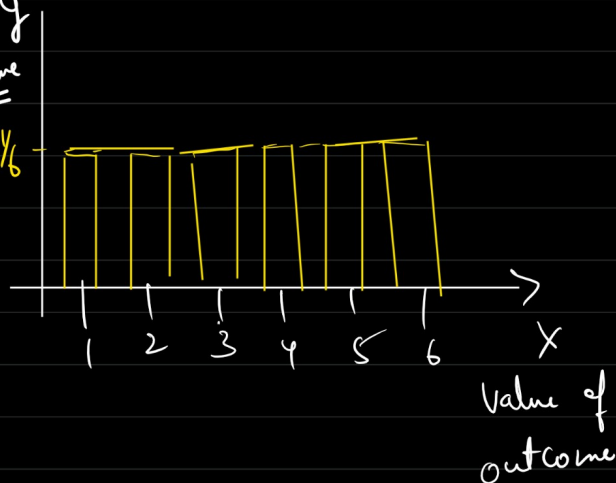
Throwing a dice \rightarrow

$\frac{1}{n}$

$$\begin{aligned} P(1) &= \frac{1}{6} \\ P(2) &= \frac{1}{6} \\ P(3) &= \frac{1}{6} \\ P(4) &= \frac{1}{6} \\ P(5) &= \frac{1}{6} \\ P(6) &= \frac{1}{6} \end{aligned}$$

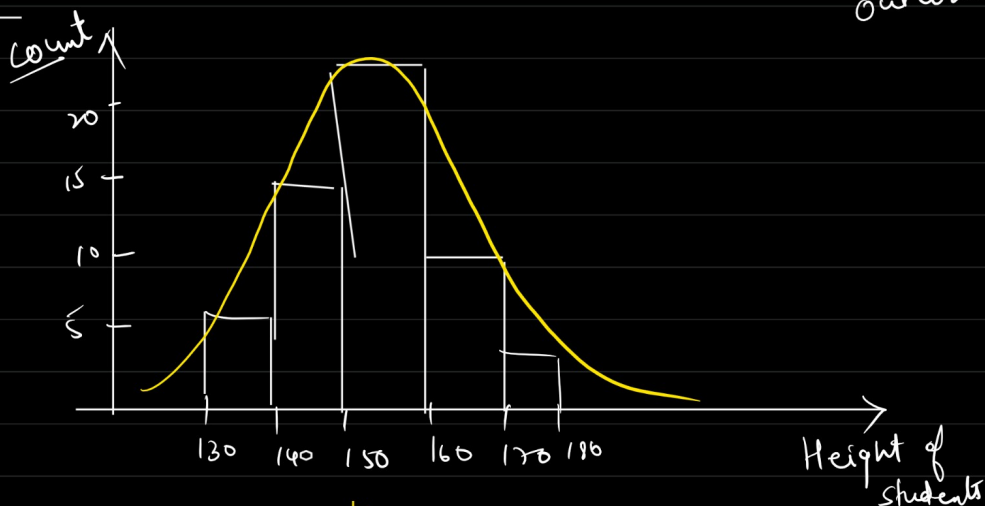
Prob
value

$\Rightarrow \frac{1}{6}$



* heights of students

150
160
160
...
...



Two types of experiment

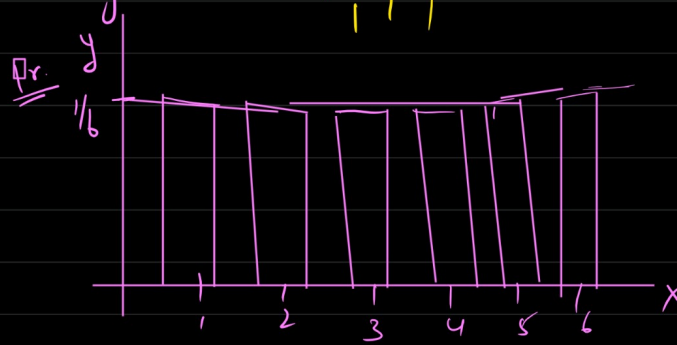
① discrete
 \Downarrow
Prob mass fn

② continuous.
 \Downarrow
Prob density function

\rightarrow Prob distribution functions

① Prob mass function (pmf) → Distribution of discrete random variable.

Example - Rolling a dice {1, 2, 3, 4, 5, 6}



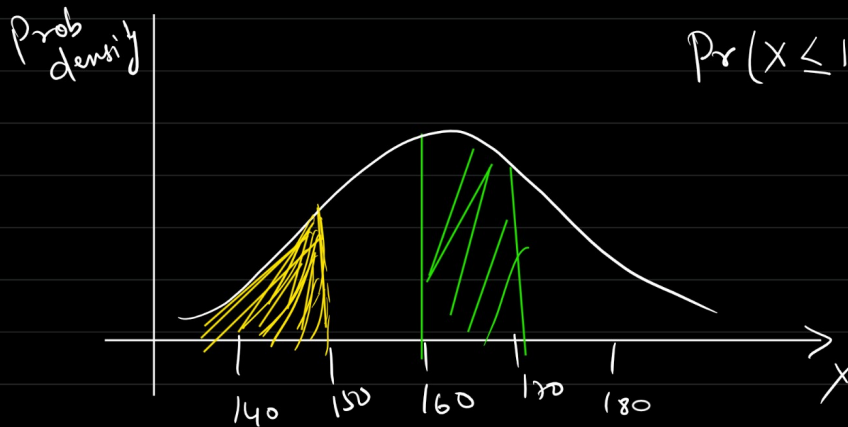
$$\begin{aligned} P(X \leq 3) &\Rightarrow \\ &= P(X=1) + P(X=2) + P(X=3) \\ &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} \end{aligned}$$

$$\begin{aligned} P(X \leq 6) &= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} \\ &= \frac{6}{6} = 1 \end{aligned}$$

Prob is between 0 & 1

② Probability density function (pdf)

→ distribution of continuous data



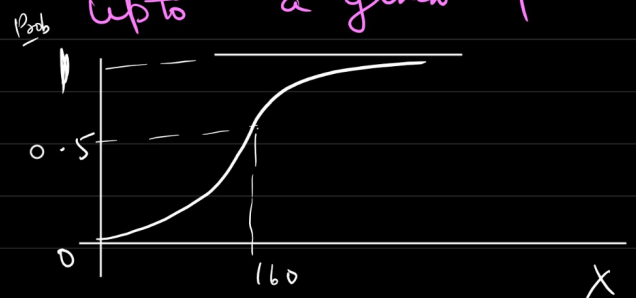
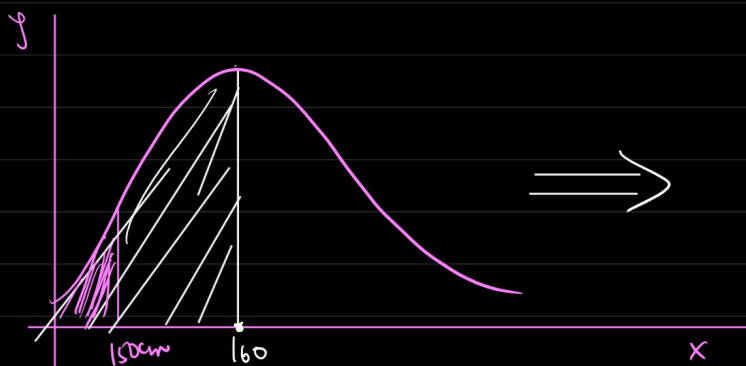
$$P(X \leq 150) = \text{Area under Curve}$$

→ Area under curve talks about probability
→ AUC — [0 to 1]

③ Cumulative distribution function (cdf)

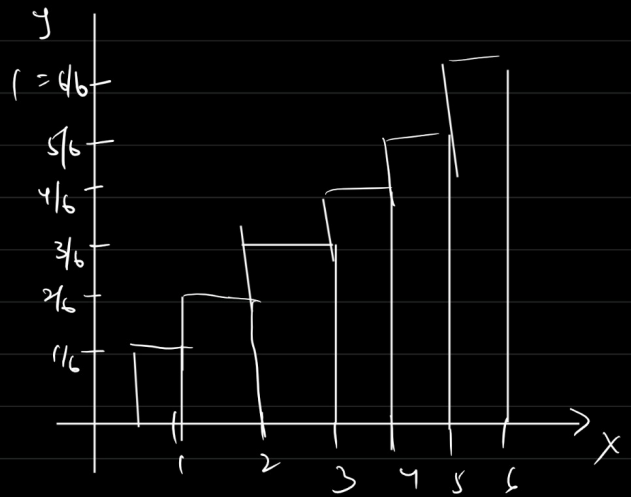
for both

→ cdf is summation of all probabilities possible upto a given point.



for pmf

x	$P(x)$	cdf
1	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$



Prob distribution function

