Random Variable

Random experiment

Sample space = Set of possible
Outcomes of Random
experiment.

Motivation :- We are not interested in the whole experiment, but value of a numerical quantity which comes out of random experiment.

Random Variable :- A function which takes input from sample space and map it to real numbers.

or Real Valued function defined on sample space.

Random experiment.

E.g.> Rolling a dice twice.

Q 1 What is the probability that sum is 5.

Q 2 What is the probability that the smaller of the outcomes is 3.

$S = \{(1,1), (1,2), \ldots, (1,6), (2,1), \ldots (2,6)$

$\ldots (6,1), \ldots (6,6)\}$

X — Sum of outcomes of two rolls.

$X : S \to \mathbb{R}$

$X \in \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

Y = Minimum of Outcome of a roll.

$Y \in \{1, 2, 3, 4, 5, 6\}$

X	Events	Probability
2	$\{(1,1)\}$	$\frac{1}{36}$
3	$\{(1,2),(2,1)\}$	$\frac{2}{36}$
4	$\{(1,3),(2,2),(3,1)\}$	$\frac{3}{36}$
5		$\frac{4}{36}$
$\vdots$		
12		$\frac{1}{36}$

X	P	X	P
2	$\frac{1}{36}$	9	$\frac{4}{36}$
3	$\frac{2}{36}$	10	$\frac{3}{36}$
4	$\frac{3}{36}$	11	$\frac{2}{36}$
5	$\frac{4}{36}$	12	$\frac{1}{36}$
6	$\frac{5}{36}$		
7	$\frac{6}{36}$		
8	$\frac{5}{36}$		

Y	Events	Probability
1	$\{(1,1), \ldots (1,6), (2,1), (3,1), (4,1), (5,1), (6,1)\}$	$\frac{11}{36}$
2	$\{(2,2), \ldots (2,6), (3,2), (4,2), (5,2), (6,2)\}$	$\frac{9}{36}$
3		$\frac{7}{36}$
4		$\frac{5}{36}$
5		$\frac{3}{36}$
6		$\frac{1}{36}$

Random experiment.

E.g.> Tossing a Coin three times.

$S = \{HHH, HHT, HTH, HTT, THT, THH, TTH, TTT\}$.

Q1 Of three tosses, how many tosses will be heads?

Q2. Of three tosses, which toss results in a head first?

X — No. of heads that appears.

	X	Y
HHH	3	1
HHT	2	1
HTH	2	1
HTT	1	1
THH	2	2
THT	1	2
TTH	1	3
TTT	0	Nil

$P(X=0) = \frac{1}{8}$

$P(X=1) = \frac{3}{8}$

$P(X=2) = \frac{3}{8}$

$P(X=3) = \frac{1}{8}$

$P(Y=1) = \frac{4}{8}$

$P(Y=2) = \frac{2}{8}$

$P(Y=3) = \frac{1}{8}$

$X \in \{0, 1, 2, 3\}$

Y — Toss in which head appears first.

$Y \in \{1, 2, 3\}$.

Discrete random Variable :→ A random Variable in Called discrete random Variable if it takes only Countably finite or infinite Values.

Apartment Complex data

- There are 12 apartments in a apartment Complex.

- Each floor has 3 apartments:
  One bedroom, two bedroom, three bedroom.

Apartment No.	Floor No.	No. of bedrooms.	Size of apartment	Distance of apartment from lift.
1	1	1	900.23	500.23
2	1	2	1175.11	325.3
3	1	3	1785.09	⋮
4	2	1	900.12	⋮
6	2	2	⋮	
7	2	3	⋮	
8	3	1		
9	3	2		
10	3	3		
11	4	1		
12	4	2	1786.30	455.37
	4	3		

[900, 1800]

Random experiment – Choosing random apartment out of 12 apartments.

X — Floor no. — $\{1, 2, 3, 4\}$

Y — No. of bedrooms – $\{1, 2, 3\}$

Z – Size of apartment – $[900, 1800]$

W – distance of apartment from lift $[300, 50]$

Discrete random Variables.

1. Maximum from a set of finite numbers.
2. outcome of rolling a die.
3. No. of students in class.
4. No. of spelling mistakes in a paragraph.

Continuous random Variable

(1) Area of apartment.

(2) Height and Weight of students.

(3) Temperature in a room.

Continuous Random Variable :-

A random Variable which takes Values in an interval.

—

Discrete RV.

Probability mass function :- Let X be a random variables which takes values $x_1, x_2, \ldots, x_N$, then Probability mass function (pmf) is defined as

$$\boxed{p(x_i) = P(X = x_i)}$$

Properties of pmf.

① $p(x_i) \geq 0$

② $\sum_{i=1}^{N} p(x_i) = 1$.

E.g. X takes values

X	0	1	2
$p(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$\Downarrow$

Is it pmf?

Yes.

E.g.1)

	x_1	x_2	x_3	x_4	x_5
X	1	2	3	4	5
$p(x_i)$	0.4	0.1	0.2	0.1	0.3

— Not a pmf

E.g.2)

X	1	2	3	4	5
$p(x_i)$	0.2	0.3	0.4	-0.1	0.2

< 0

Not a pmf.

Example. X $\quad x_0 \quad x_1 \quad x_2 \quad x_3 \quad \cdots$

$\qquad\quad 0 \quad 1 \quad 2 \quad 3 \quad \cdots$ $\qquad \lambda > 0$

$$p(x_i) - P(X = x_i) \quad \frac{c\lambda^0}{0!} \quad \frac{c\lambda^1}{1!} \quad \frac{c\lambda^2}{2!} \quad \frac{c\lambda^3}{3!} \cdots$$

Can you find c such that $p(x_i)$ is pmf.

(1) $p(x_i) \geq 0 \implies c \geq 0.$

$$c\frac{\lambda^0}{0!} + c\frac{\lambda^1}{1!} + c\frac{\lambda^2}{2!} + \dots = 1$$

$$\Rightarrow c\left(1 + \frac{\lambda}{1!} + \frac{\lambda^2}{2!} + \dots\right) = 1$$

$$e^{\lambda} = 1 + \lambda + \frac{\lambda^2}{2!} + \dots$$

$$\Rightarrow c\, e^{\lambda} = 1$$

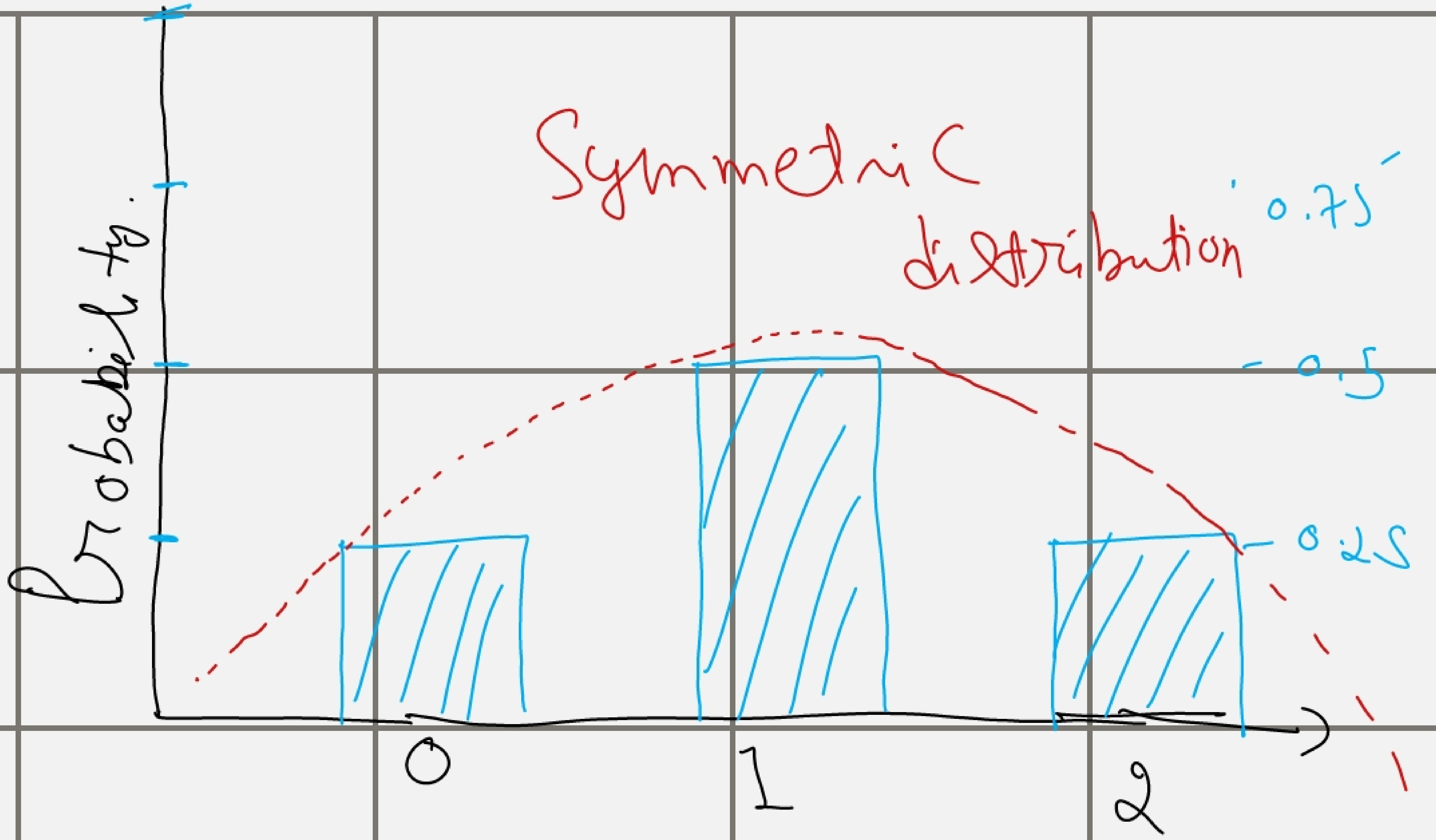
$$\Rightarrow \boxed{c = e^{-\lambda}}$$

$$p(x_i) = P(X = x_i) = c \cdot \frac{\lambda^i}{i!}$$

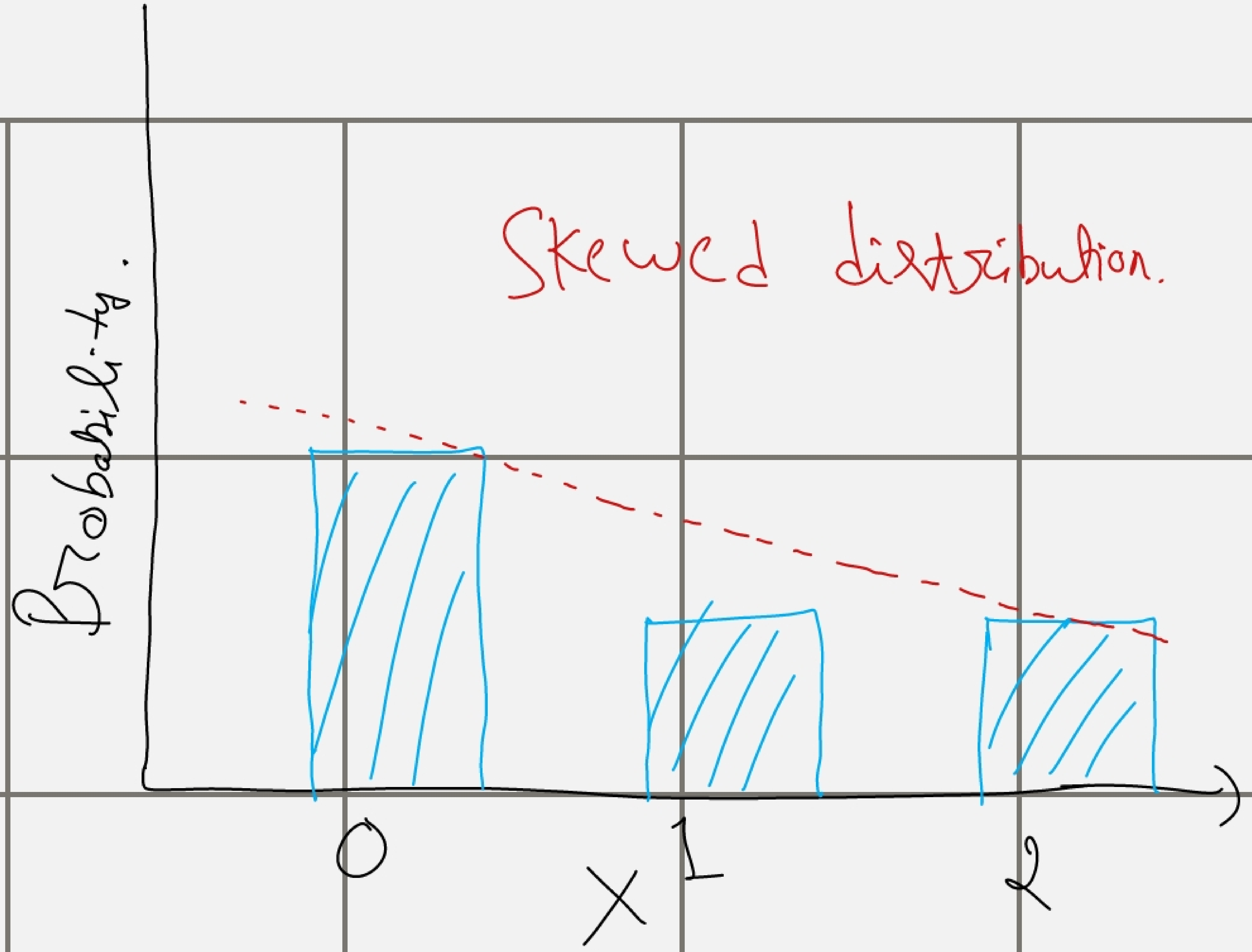
$$\boxed{p(x_i) = e^{-\lambda} \frac{\lambda^i}{i!}}$$

Graph of probability mass function

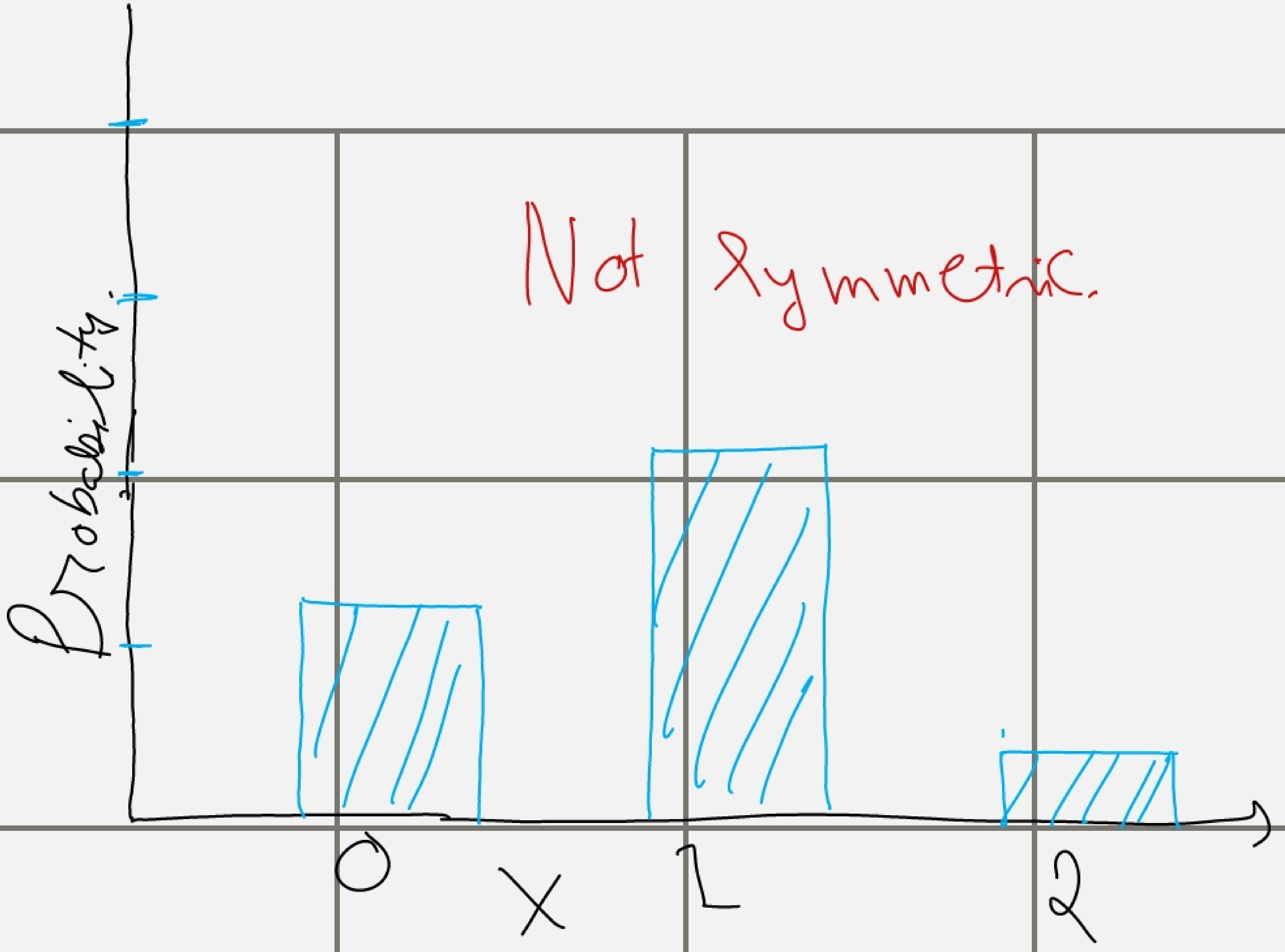
X	0	1	2
$P(X=x_i)$	0.25	0.5	0.25



X	0	1	2
$P(X=x_i)$	0.5	0.25	0.25



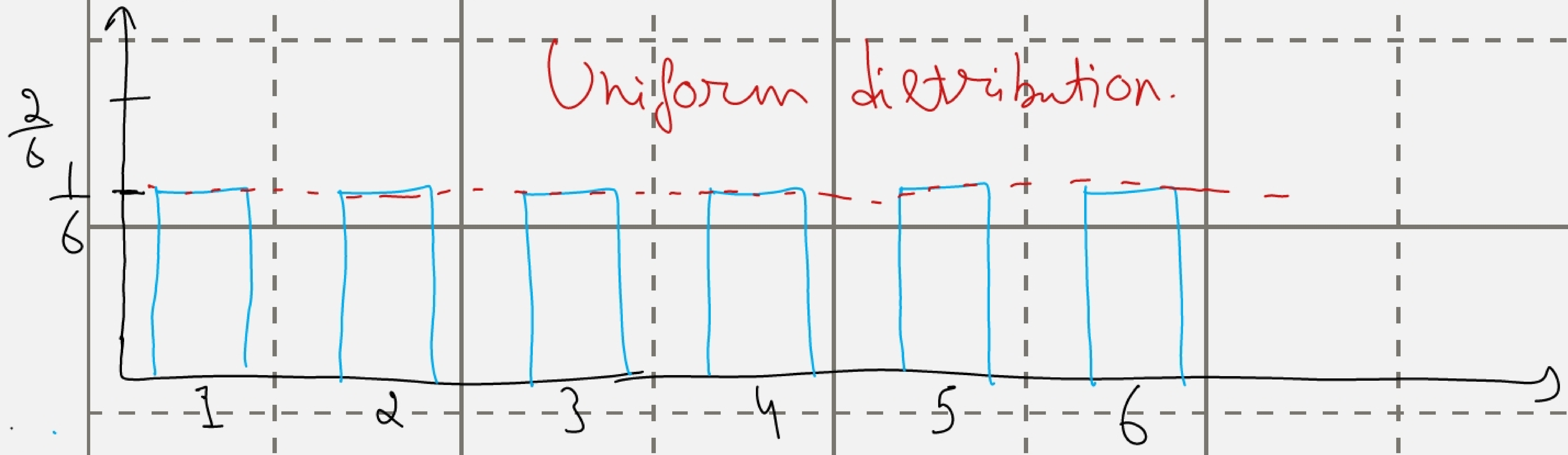
X	0	1	2
$P(X=x_i)$	0.35	0.55	0.10



Not Symmetric.

E.g. Rolling a die

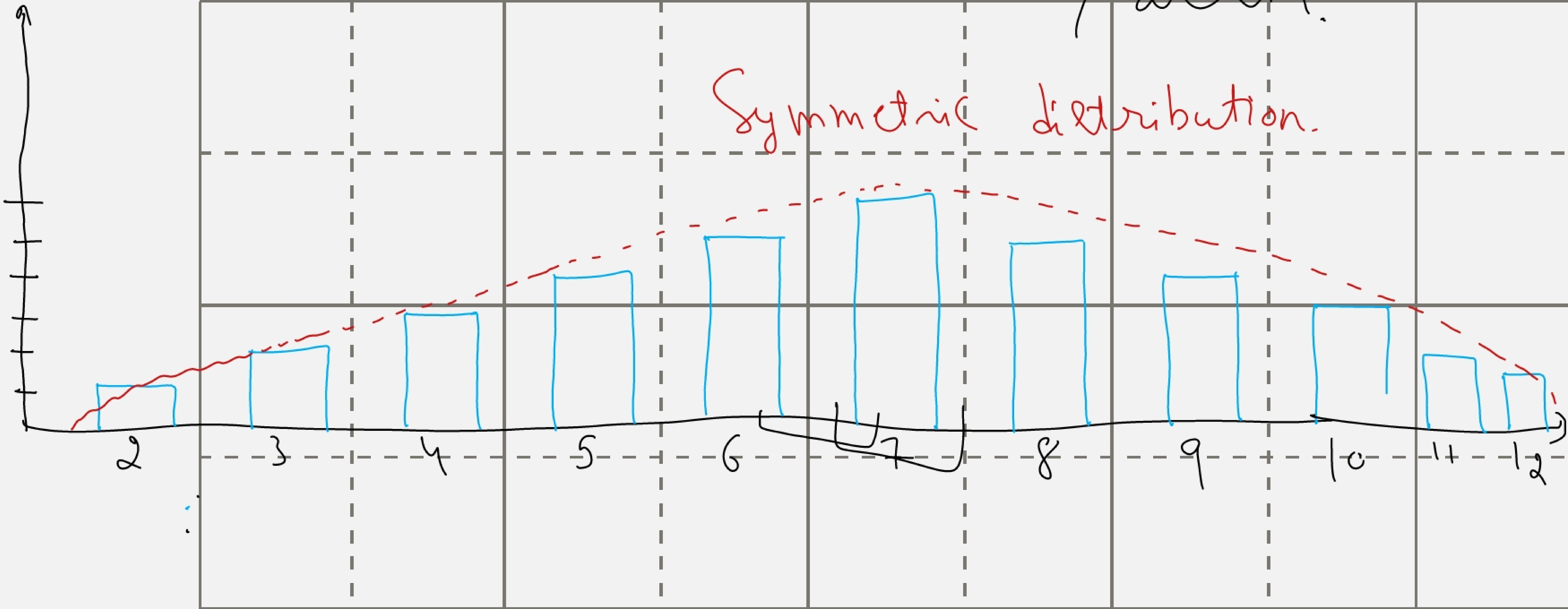
X	1	2	3	4	5	6
$P(X=x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$.



$E[X] = 7$

X	2	3	4	5	6	7	8	9	10	11	12
$P(X=x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Try it for Y also?



pmf →

X

$P(X=x_i)$

$\|$

$p(x_i)$

$P(X \le a)$, $P(X \ge a)$

$P(X \ge 5)$ ✓

$$\boxed{F(a) = P(X \le a)}$$

$F : \mathbb{R} \to [0,1]$

Cdf.

Cumulative distribution function:

$F : \mathbb{R} \to [0,1]$ defined by

$$\boxed{F(a) = P(X \le a)} \text{ — Cdf.}$$

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

$a = -1$

$P(X \le 0)$

E.g>

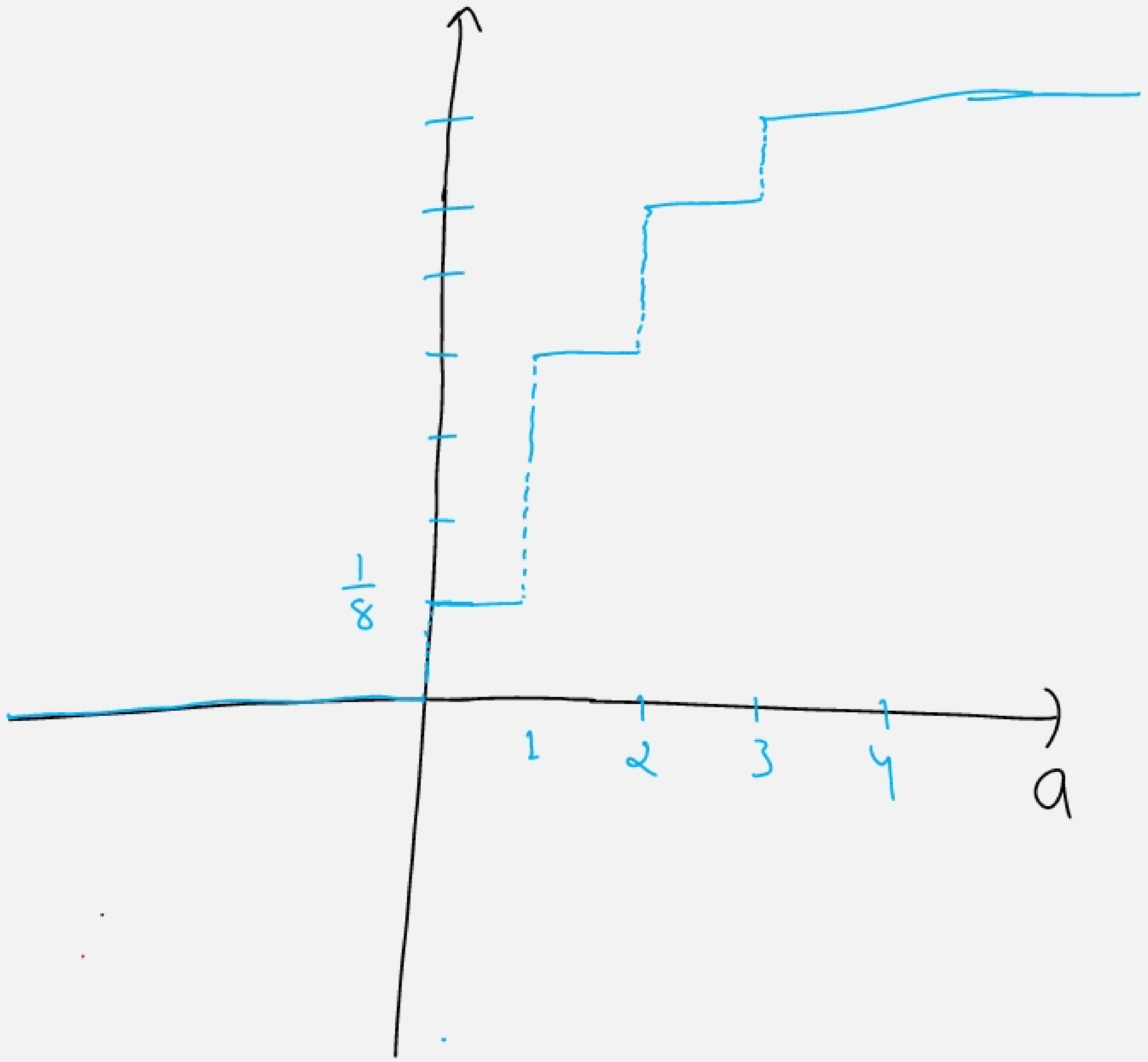
Tossing a Coin thrice

X – No. of heads in an outcome

X	0	1	2	3
$P(X=x_i)=$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$F(a) = \begin{cases} 0 & a < 0 \\ \frac{1}{8} & 0 \le a < 1 \\ \frac{4}{8} & 1 \le a < 2 \\ \frac{7}{8} & 2 \le a < 3 \\ 1 & a \ge 3 \end{cases}$$

$$F(a) = \begin{cases} 0, & a < 0 \\ \frac{1}{8}, & 0 \le a < 1 \\ \frac{4}{8} & 1 \le a < 2 \\ \frac{7}{8} & 2 \le a < 3 \\ 1 & a \ge 3 \end{cases}$$



<u>Remark:</u> Let X be a discrete random variable taking values $x_1, x_2, \dots, x_n$, then cdf $F(a)$ will be a step function

Case Study :-

To analyze no. of Credit Cards owned by a population.

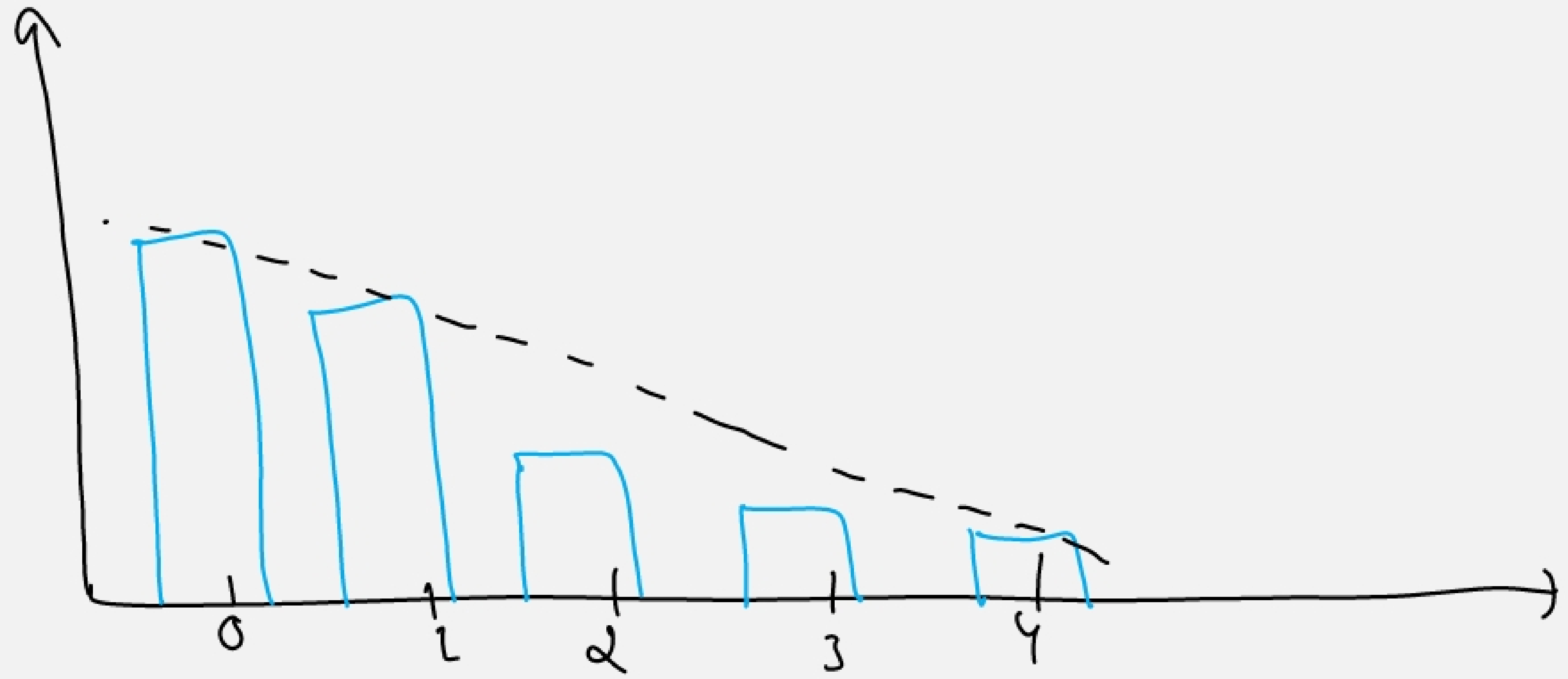
- Collect data :- Ask people how many Credit Cards they own.

$S = \{x_1, x_2, \ldots, x_n\}$

Random experiment :- Picking a random individual from the population.

X — No. of Credit Cards owned by a person.

X	0	1	2	3	4
$P(X = x_i)$	0.42	0.36	0.14	0.06	0.02



Question:- Choose an individual at random.

Is he/she more likely to have 0 Credit Cards or 2 or more Credit Cards?

pmf $\leftarrow P(X=0) = 0.42 = 42\%$

Cdf $\longrightarrow P(X \geq 2) = 0.14 + 0.06 + 0.02 = 0.22 = 22\%$

$1 - P(X<2) = 0.22$

Question:- Random sample of 1000 individuals, we ask those 1000 individuals, how many Credit Cards they own?

- Everyone is saying, they own a credit cards

  - There is some problem in sample.

(2) 138 people respond that they own two credits cards.

$\boxed{1000 \times 0.14 = 140}$

Choose an individual at random, how many Credit Cards would you expect that individual to own?

0

→ This we will answer with expected Value.

Example :- let us see one game.

Rolling a single die

pmf.	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$
X	1	2	3	4	5	6
	$+1$	-2	$+3$	-4	$+5$	-6

Roll die $\overset{n}{\overset{\shortparallel}{100}}$ times

X	Wining Y	Freuency.		
1	+1	18	0.18	$\frac{1}{6}$
2	−2	12	0.12	$\frac{1}{6}$
3	+3	20	0.20	$\frac{1}{6}$
4	−4	15	0.15	$\frac{1}{6}$
5	+5	15	0.15	$\frac{1}{6}$
6	−6	20	0.20	$\frac{1}{6}$
		100	1	

$$\frac{1}{6}-\frac{2}{6}+\frac{3}{6}-\frac{4}{6}+\frac{5}{6}-\frac{6}{6} = \frac{9-12}{6} = \frac{-3}{6}$$

Average winning $= -\frac{1}{2} = -0.5$

$$1\times0.18 - 2\times0.12 + 3\times0.20 - 4\times0.15 + 5\times0.15 - 6\times0.20 = -0.51$$

$X \sim x_1, x_2, \ldots, x_n$

$$\boxed{E[X] = \sum_{i=1}^{n} \underline{x_i}\, \underline{P(X=x_i)}}$$

Expectation of a random variable :- Let X be a discrete random variable taking values $x_1, x_2, \ldots, x_n$, then expected value of X is defined as

$$E[X] = \sum_{i=1}^{n} x_i\, \underline{P(X=x_i)}.$$

Example:- A player pays $2 to roll a fair six sided die. The payout rules are

Roll a 1 — Win $5

Roll a 6 — Win $3

Any other No. — $0.

$x_1 \quad x_2 \ldots x_6$

Amount						
X —	1	2	3	4	5	6
Y	5	0	0	0	0	3
Prob.	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

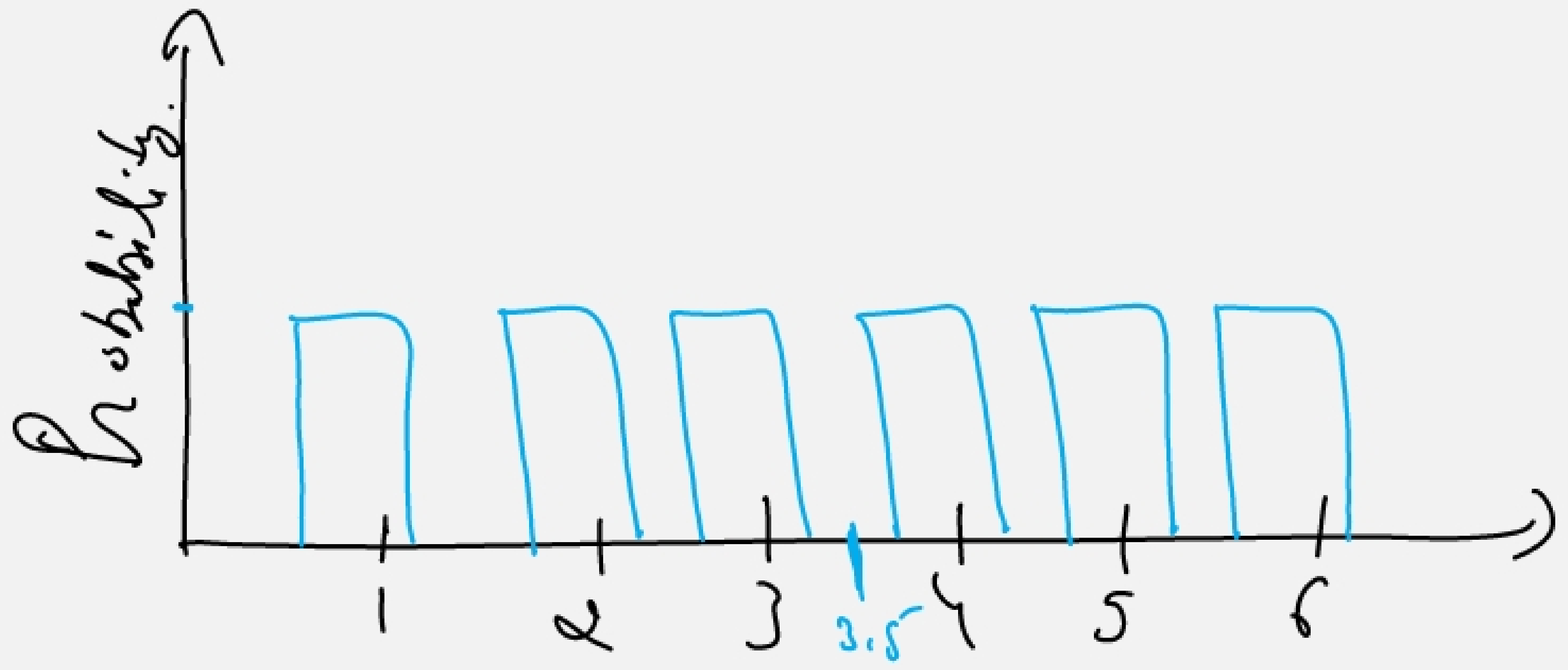
X — Outcome of die

$$E[Y] = 5\times\frac{1}{6} + 3\times\frac{1}{6} = \frac{8}{6} = \frac{4}{3} = 1.33$$

We are also paying $2.

Profit = 1.33 − 2 = −0.67

$$E[X] = 1\times\frac{1}{6} + 2\times\frac{1}{6} + 3\times\frac{1}{6} + 4\times\frac{1}{6} + 5\times\frac{1}{6} + 6\times\frac{1}{6} = \frac{7}{2} = 3.5$$



E.g.> Rolling a die twice

X – Sum of outcomes.

X	2	3	4	5	6	7	8	9	10	11	12
$P(X=x_i)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$E[X] = 2\times\frac{1}{36} + 3\times\frac{2}{36} + \dots + 12\times\frac{1}{36} = 7.$$

Example :- Tossing a Coin thrice.

X — No. of heads in the tosses

X	0	1	2	3
$P(X=x_i)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

$$E[X] = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2} = 1.5$$

Example. → Bernoulli Random Variable.

X —	0	1
$P(X=x_i)$	$1-p$	p

$$\boxed{0 \leq p \leq 1}$$

$$E[X] = p$$

Propatics of expectation.

(1) X — Random Variable

aX — Random Vaiable

$g(X)$

a in a real No.

$aX + b$ — Random Vaiable.

X	x_1	x_2	...	x_n
pmf	$p(x_1)$	$p(x_2)$	...	$p(x_n)$

$$\sum_{i=1}^{n} p(x_i) = 1$$

aX	ax_1	ax_2	...	ax_n
pmf	$p(x_1)$	$p(x_2)$	...	$p(x_n)$

$$\boxed{E[aX+b] = aE[X]+b}$$

$$= \sum_{i=1}^{n} (ax_i + b)\, p(x_i)$$

$$= \sum_{i=1}^{n} ax_i\, p(x_i) + b \sum_{i=1}^{n} p(x_i)$$

$$= aE[X] + \underline{b}$$

$$E[\underline{X^2}] = (\underline{E[X]})^2$$

$$\sum_{i=1}^{n} x_i^2 p(x_i) \neq \left(\sum_{i=1}^{n} x_i p(x_i)\right)^2$$

Think of a

Counter example.

Example: Bulk purchase discount.

A store offers a discount on bulk purchases | 150 - Cost of each item.

Buy 1-4 items : No discount
5-9 items : 10% discount
10 or more : 20% discount

X	3	6	10
pmf	0.3	0.4	0.3
Y	150	270	400

X - No. of items a customer buys.

$P(X=3)=0.3$, $P(X=6)=0.4$, $P(X=10)=0.3$

Y = Total Cost after discount.

$E[Y]$

Recall :- Expectation of a random Variable.

$$E[X] = \sum_{i=1}^{n} x_i \, p(x_i).$$

$X = 0$ with probability 1

$$Y = \begin{cases} -1 & \text{with probability } \frac{1}{2} \\ 1 & \text{'' '' } \frac{1}{2} \end{cases}$$

$$Z = \begin{cases} -100 & \text{'' '' } \frac{1}{2} \\ 100 & \text{'' '' } \frac{1}{2} \end{cases}$$

$$E[X] = E[Y] = E[Z] = 0$$

X	0
$p(x)$	1

Y	-1	1
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

Z	-100	100
$p(x)$	$\frac{1}{2}$	$\frac{1}{2}$

$E[X^2] = 0$

$Var(X) = 0$

$E[Y^2] = 1$

$Var(Y) = 1$

$E[Z^2]$

$= \frac{100^2}{2} + \frac{100^2}{2}$

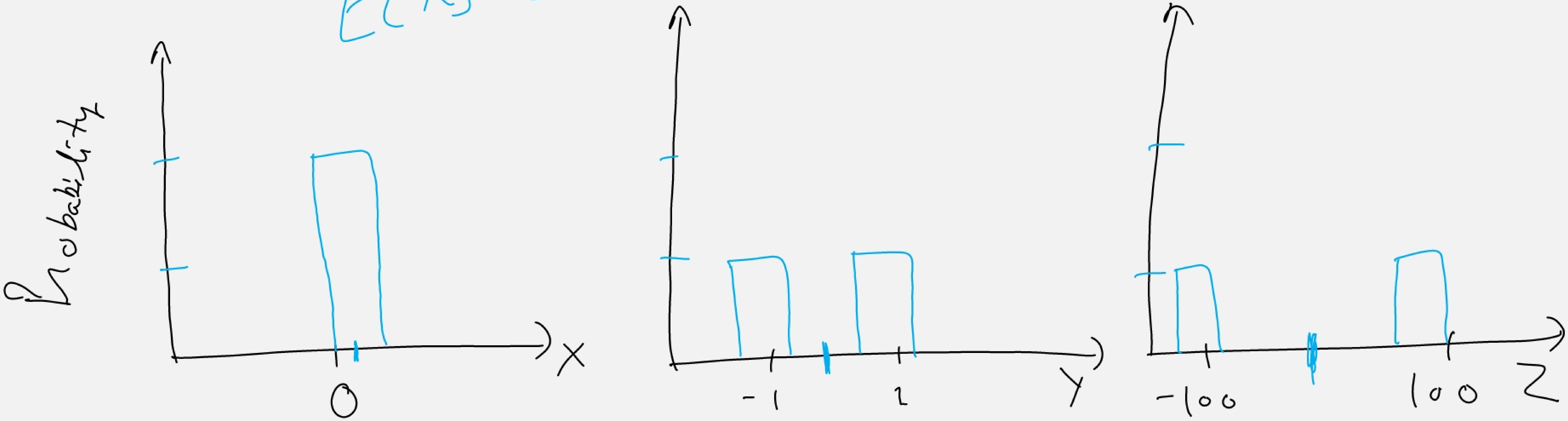
$= 100^2$

$= 10{,}000$

$Var(Z)$

$= 10{,}000.$

$E[X] = E[Y] = E[Z] = 0$



$$E\left[\,\underline{|X-\mu|}\,\right]$$

$$E\left[\underline{(X-\mu)^2}\right]$$

Let X be a discrete random Variable, then Variance of X is defined as

$$Var(X) = E\left[(X-\mu)^2\right]$$

$$\boxed{\mu = E[X]}$$

Let X be a random Variable taking Values $x_1, x_2, \ldots, x_n$, then

$$Var(X) = E[(X-\mu)^2] = \sum_{i=1}^{n} (x_i-\mu)^2 p(x_i)$$

$$= \sum_{i=1}^{n} (x_i^2 + \mu^2 - 2\mu x_i)\, p(x_i)$$

$$= \sum_{i=1}^{n} x_i^2 p(x_i) + \mu^2 \sum_{i=1}^{n} p(x_i) - 2\mu \sum_{i=1}^{n} x_i p(x_i)$$

$$\boxed{E[g(X)] = \sum_{i=1}^{n} g(x_i)\, p(x_i)}$$

$$Var(X) = E[X^2] + \mu^2 - 2\mu E[X]$$

$$\underline{Var(X)} = E[X^2] - \mu^2 = E[X^2] - (E[X])^2$$

Example: X — Outcome of a die

X — $\{1, 2, 3, 4, 5, 6\}$

$E[X] = \frac{7}{2}$

$E[X^2] = \frac{91}{6}$

X	1	2	3	4	5	6
$p(x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

$$Var(X) = E[X^2] - (E[X])^2 = \frac{91}{6} - \frac{49}{4} = \frac{35}{12} = 2.9$$

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2$$

$$\mathrm{Var}[X],\ \mathrm{Var}(X)$$

$$E[aX+b] = aE[X] + b$$

$$\mathrm{Var}(aX+b) = E[(aX+b)^2] - \left(E[(aX+b)]\right)^2$$

$$= E[a^2X^2 + 2abX + b^2] - (aE[X]+b)^2$$

$$= a^2E[X^2] + 2\cancel{ab}E[X] + \cancel{b^2} - \left(a^2(E[X])^2 + 2\cancel{ab}E[X] + \cancel{b^2}\right)$$

$$= a^2\,\mathrm{Var}(X)$$

$$\mathrm{Var}(aX+b) = a^2\,\mathrm{Var}(X)$$

Standard deviation of a random Variable

$$SD(X) = \sqrt{Var(X)}\ ,\ \text{+ve square root.}$$

Geometric distribution.

1.> Let X be a Bernoulli trial. The outcome of trial can be "success" or "failure".

2.> The trials are identical and independent. that means outcome of one trial will not affect the outcome of other trial.

3.> Random Variable X denotes the no. of trials required to obtain first success.

Sample space for random experiment will be

$S = \{s, fs, ffs, fffs, ffffs, \dots\}$ $\quad 0<p<1$

$|r|<1$

X — $1, 2, 3, 4, \dots$

discrete random variable.

Probability of success is p

" " failure is $q = 1-p$

$\sum_{n=1}^{\infty} r^n = \frac{1}{1-r}$

X	1	2	3	4	5	...
$p(x_i) = P(X=x_i)$	p	$(1-p)p$	$(1-p)^2 p$	$(1-p)^3 p$	$(1-p)^4 p$	

$p(i)$

$$P(X=n) = (1-p)^{n-1} p$$
$$= q^{n-1} p$$

$$\sum_{n=1}^{\infty} P(X=n) = p\sum_{n=1}^{\infty} q^{n-1} = p\,\frac{1}{1-q} = \frac{p}{p} = 1$$