

## Probability

We all learned simple probability before!

e.g. Roll a die  $\rightsquigarrow$    $P(2) = \frac{1}{6}$

Toss a coin  $\rightsquigarrow$    $P(\text{head}) = \frac{1}{2}$

Goal: Make things formal

Discrete  $\rightarrow$  Continuous

## Terminologies :

- 1) **Possible outcome** : possible result of a single experiment
- 2) **Sample space** : collection of all possible outcomes
- 3) **Events** : Any collection of possible outcomes

e.g. A box containing 3 balls, marked 1, 2 and 3

Randomly select one.

1) Possible outcome 1, 2, 3

2) Sample space  $S = \{1, 2, 3\}$

3) Events  $\phi, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}$

↑  
empty

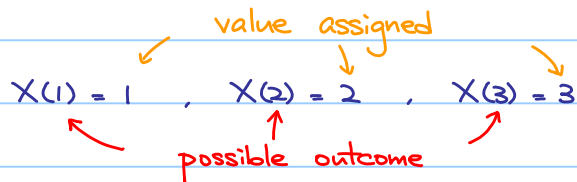
Event A = the number marked on the selected ball is less than 3 =  $\{1, 2\}$

Event B = the number marked on the selected ball is 0 =  $\phi$

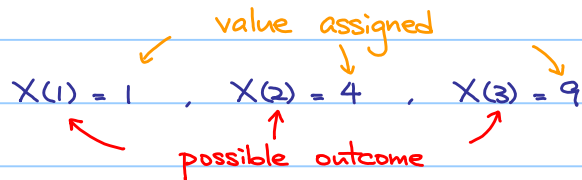
4) **Random variable** : A function  $X$  that assigns a real number to each possible outcome  
(i.e.  $X: S \rightarrow \mathbb{R}$ )

Refer to the previous example :

(i) Let  $X$  be the number marked on the selected ball.



(ii) Let  $X$  be the square of the number marked on the selected ball.



If the image of  $X$  is finite or countably infinite, then  $X$  is called a discrete random variable, if the image of  $X$  is uncountably infinite, then  $X$  is called a continuous random variable.

e.g. Toss a coin 3 times

$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$  8 possible outcomes

Let  $X$  be the number of times a head comes out.

$X(HHH) = 3$ ,  $X(HHT) = 2$ , ...

(image of  $X = \{0, 1, 2, 3\}$  which is finite, so  $X$  is discrete.)

We can use  $X=2$  to denote all possible outcomes with value 2

$X=2 = \{HHT, HTH, THH\}$

so  $X=2$  can be regarded as an event.

$X=2$  is the event of getting 2 heads.

In general,  $X=x$  is the event of getting  $x$  heads.

e.g. Toss a coin until a head comes out.

$$S = \{H, TH, TTH, TTTH, \dots\}$$

Let  $X$  be the number of trials.

$$X(H) = 1, X(TH) = 2, X(TTH) = 3, \dots$$

(image of  $X = \{1, 2, 3, \dots\}$  which is countably infinite, so  $X$  is discrete.)

e.g. There are four students  $A, B, C, D$  with heights  
160 cm, 165 cm, 170 cm, 180 cm respectively.

Randomly pick two of them.

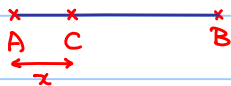
$$S = \{AB, AC, AD, BC, BD, CD\}$$

Let  $X$  be the sum of the heights of the two selected students

$$X(AB) = 160 + 165 = 325, \dots$$

(image of  $X$  is finite, so  $X$  is discrete.)

e.g. Suppose AB is a rope of length 10cm and AB is cut into two pieces.



$$S = (0, 10)$$

Let  $X$  be the square of length of  $AC^2$ , so  $X(x) = x^2$ .

(image of  $X = (0, 100)$  which is uncountably infinite, so  $X$  is continuous.)



## Probability Distribution (Discrete random variable)

$S$ : sample space

$X$ : discrete random variable

A probability distribution (or probability density function pdf) is a function  $p$  such that

$$p(x) = P(X=x)$$

= Probability of the event  $X=x$

e.g. Toss a coin 3 times

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

Let  $X$  be the number of times a head comes out.

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

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

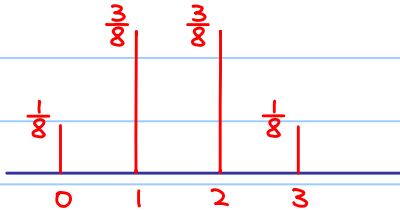
$$p(2) = P(X=2) = \frac{3}{8}$$

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

$$p(x) = 0 \quad \text{if } x \neq 0, 1, 2, 3$$

Note :  $\sum_{x; p(x) \neq 0} p(x) = 1$  ( summing over  $x$  with  $p(x) \neq 0$  )

Graphical representation :



e.g. Toss a coin until a head comes out.

$$S = \{H, TH, TTH, TTTH, \dots\}$$

Let  $X$  be the number of trials.

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

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

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

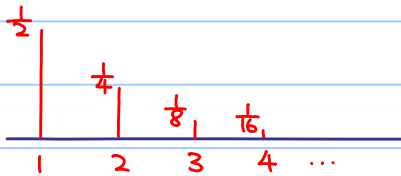
$$p(4) = P(X=4) = \frac{1}{16}$$

$\vdots$

$$\therefore p(x) = \begin{cases} \frac{1}{2^{x+1}} & \text{if } x \text{ is a nonnegative integer} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned} \text{Again } \sum_{x: P(x) \neq 0} P(x) &= P(0) + P(1) + P(2) + P(3) + \dots \\ &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \dots \\ &= 1 \end{aligned}$$

Graphical representation :



## Probability Distribution (Discrete random variable)

$S$ : sample space

$X$ : discrete random variable

A probability distribution (or probability density function pdf) is a function  $p$  such that

1)  $p(x) = P(X=x) = \text{Probability of the event } X=x$

2)  $p(x) \geq 0$

3)  $\sum_{x: p(x) \neq 0} p(x) = 1$

## Expected Value of a Discrete Random Variable

$X$ : discrete random variable

Expected Value of  $X = E(X)$

$$= \sum_{x: p(x) \neq 0} x p(x)$$

We also call it mean and denote it by  $\mu$ .

e.g. Roll a die

Sample space =  $S = \{1, 2, 3, 4, 5, 6\}$

Let  $X$  be the random variable that denotes the number facing up.

$$p(x) = \begin{cases} \frac{1}{6} & \text{if } x = 1, 2, 3, 4, 5, 6 \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{aligned} E(X) &= \sum_{x: p(x) \neq 0} x p(x) \\ &= 1 \cdot p(1) + 2 \cdot p(2) + \dots + 6 \cdot p(6) \\ &= 3.5 \end{aligned}$$



Q: What is the meaning of this 3.5?

Think: If we roll a die for 600 times,

we expect (but NOT necessary to be true) that each number occurs  $\frac{1}{6} \times 600 = 100$  times, so the average of the result is

$$\frac{100 \times 1 + 100 \times 2 + \dots + 100 \times 6}{600} = \frac{1}{6} \times 1 + \frac{1}{6} \times 2 + \dots + \frac{1}{6} \times 6 = 3.5$$

## Variance of a Discrete Random Variable

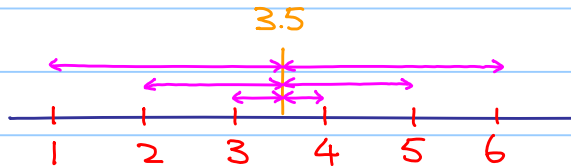
$X$ : discrete random variable

Variance of  $X = \text{Var}(X)$

$$= \sum_{x: p(x) \neq 0} (x - \mu)^2 p(x)$$

e.g. Roll a die

$$\begin{aligned} \text{Var}(X) &= (1-3.5)^2 \cdot \frac{1}{6} + (2-3.5)^2 \cdot \frac{1}{6} + (3-3.5)^2 \cdot \frac{1}{6} + \\ &\quad (4-3.5)^2 \cdot \frac{1}{6} + (5-3.5)^2 \cdot \frac{1}{6} + (6-3.5)^2 \cdot \frac{1}{6} . \\ &= \frac{35}{12} \end{aligned}$$



Variance provides a measure of the tendency of the values of  $X$  to spread out of the expected value  $E(X)$ .

standard deviation  $\sigma := \sqrt{\text{Var}(X)}$

Recall: In statistics,  $\sigma = \sqrt{\frac{\sum f_i (x_i - \mu)^2}{n}}$

$$= \sqrt{\sum p_i (x_i - \mu)^2}$$

same formula!

$x_i$ : values

$f_i$ : frequency

$n$ : total number

$$p_i = \frac{f_i}{n}$$

More examples :

## Geometric Random Variables

Experiment : Only two possible result	success (S)	$p$
	failure (F)	$1-p$

Continue to try until one success trial.

Sample space =  $\{S, FS, FFS, FFFS, \dots\}$

Let  $X$  be the random variable that counts the number of trials.

The event  $X=x$  is just the event  $\underbrace{FF \dots F}_{x-1 \text{ times}} S$   
 $\uparrow$   
 $x$ -th trial

$$p(x) = \begin{cases} (1-p)^{x-1} p & \text{if } x \text{ is a positive integer} \\ 0 & \text{otherwise} \end{cases}$$

$$\begin{aligned}
 \textcircled{1} \sum_{x; p(x) \neq 0} p(x) &= \sum_{n=1}^{\infty} (1-p)^{n-1} p \\
 &= p [1 + (1-p) + (1-p)^2 + \dots] \\
 &= p \cdot \frac{1}{1-(1-p)} \\
 &= 1
 \end{aligned}$$

$$\begin{aligned}
 \textcircled{2} E(x) &= \sum_{x; p(x) \neq 0} x p(x) \\
 &= \sum_{n=1}^{\infty} n p(n) \\
 &= \sum_{n=1}^{\infty} n p (1-p)^{n-1} \\
 &= p \sum_{n=1}^{\infty} n (1-p)^{n-1} \\
 &= \frac{1}{p}
 \end{aligned}$$

$$\begin{aligned}
 \sum_{n=1}^{\infty} x^n &= \frac{1}{1-x} \\
 \frac{d}{dx} \left( \sum_{n=1}^{\infty} x^n \right) &= \frac{d}{dx} \left( \frac{1}{1-x} \right) \\
 \sum_{n=1}^{\infty} n x^{n-1} &= \frac{1}{(1-x)^2}
 \end{aligned}$$

put  $x = 1-p$ ,  $\sum_{n=1}^{\infty} n (1-p)^{n-1} = \frac{1}{p^2}$

$$\textcircled{3} \text{Var}(X) = \sum_{x:p(x) \neq 0} (x - \mu)^2 p(x)$$

$$= \sum_{n=1}^{\infty} \left(n - \frac{1}{p}\right)^2 \cdot p(1-p)^{n-1}$$

$$= p \sum_{n=1}^{\infty} n^2 (1-p)^{n-1} - 2 \sum_{n=1}^{\infty} n(1-p)^{n-1} + \frac{1}{p} \sum_{n=1}^{\infty} (1-p)^{n-1}$$

$$= \frac{2-p}{p^2} - \frac{2}{p^2} + \frac{1}{p^2}$$

$$= \frac{1}{p^2}$$

$$\sum_{n=1}^{\infty} nx^{n-1} = \frac{1}{(1-x)^2}$$

$$\sum_{n=1}^{\infty} nx^n = \frac{x}{(1-x)^2}$$

$$\frac{d}{dx} \sum_{n=1}^{\infty} nx^n = \frac{d}{dx} \frac{x}{(1-x)^2}$$

$$\sum_{n=1}^{\infty} n^2 x^{n-1} = \frac{1+x}{(1-x)^3}$$

put  $x = (1-p)$ ,

$$\sum_{n=1}^{\infty} n^2 (1-p)^{n-1} = \frac{2-p}{p^3}$$

$$\sum_{n=1}^{\infty} n(1-p)^{n-1} = \frac{1}{p^2}$$

$$\sum_{n=1}^{\infty} (1-p)^{n-1} = \frac{1}{p}$$

e.g. Printing document

probability of jamming for printing one page = 0.05% = 0.0005

Expect value =  $\frac{1}{p} = 2000$

We expect the printer to have one jamming on 2000-th page.