

# Poisson Distribution

- Suppose we are counting the number of occurrences of an event in a given unit of time, distance, area, or volume.

For Example:

- The number of car accidents in a day
- The number of defective items in a lot

# Poisson Distribution

**Suppose:**

- Events are occurring independently
- The probability that an event occurs in a given length of time does not change through time.
- Then  $X$ , the number of events in a fixed unit of time, has a Poisson Distribution



# Poisson Distribution

- Then  $X$ , the number of events (success) in a fixed unit of time, has a Poisson Distribution

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}; x = 0, 1, \dots$$

Where,  $\lambda$  is the expected number of occurrences, which is  $np$

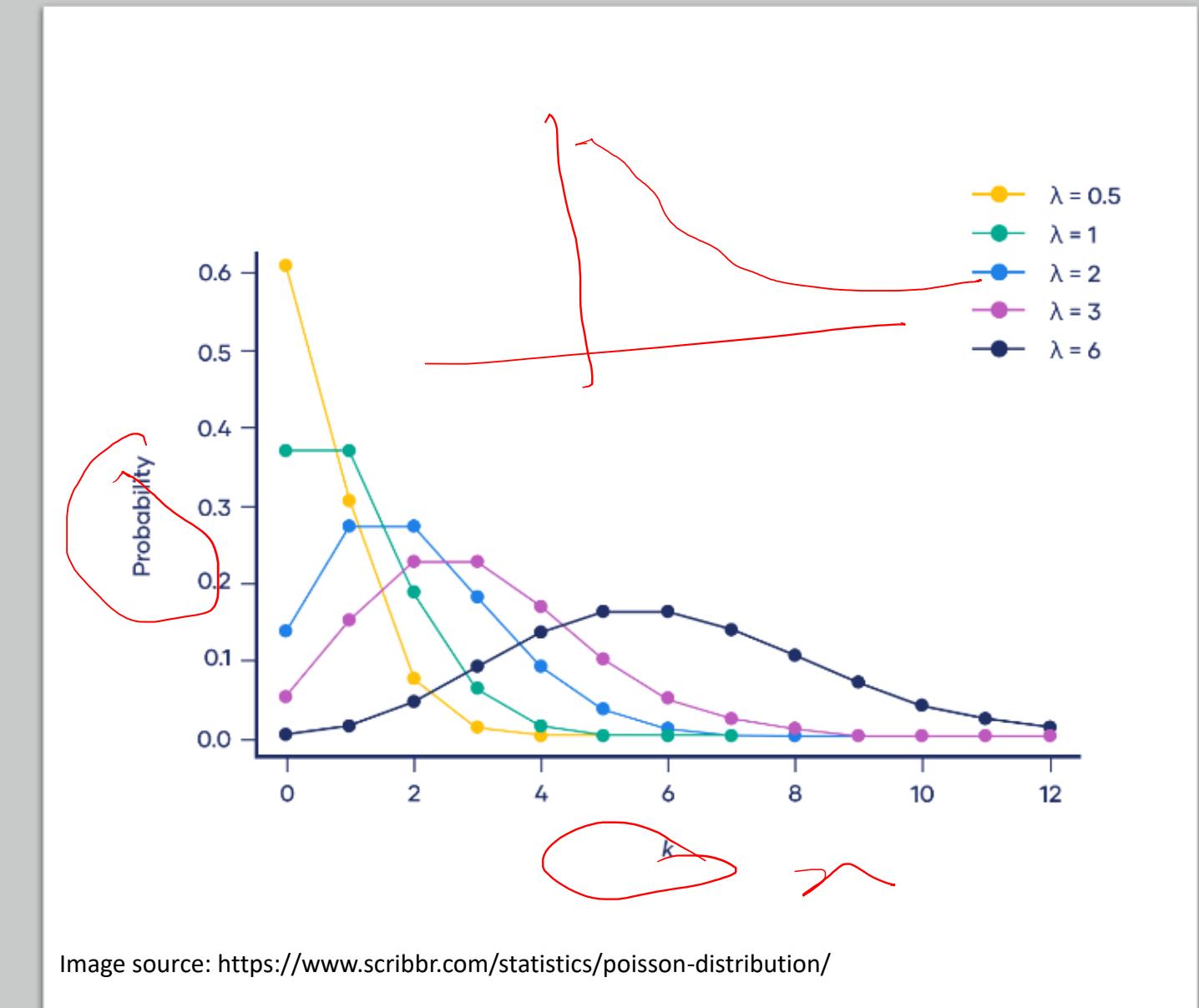
$x$ : number of successes

$e$ : a constant equal to approximately 2.71828



# Poisson Distribution

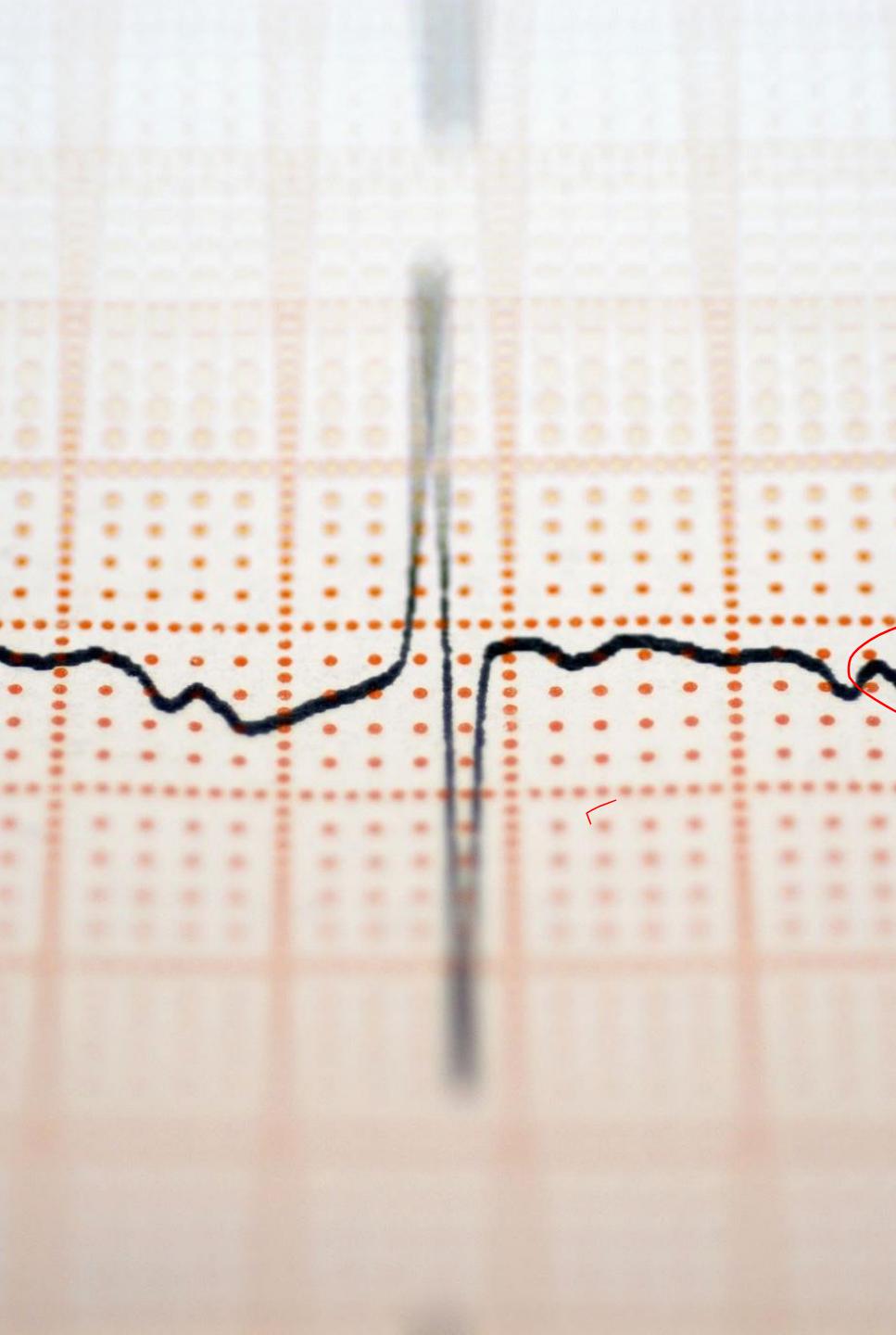
- A Poisson distribution is a discrete probability distribution.
- It gives the probability of an event happening a certain number of times ( $x$  or  $k$ ) within a given interval of time or space.
- The Poisson distribution has only one parameter,  $\lambda$  (lambda), which is the mean number of events.



# Binomial vs Poisson Distribution

- Binomial distribution describes the distribution of binary data from **a finite sample**. Thus it gives the probability of getting r events out of n trials.  

- Poisson distribution describes the distribution of binary data from **an infinite sample (or very large sample)**. Thus it gives the probability of getting r events in a population.



# Mean and Variance of Poisson Distribution

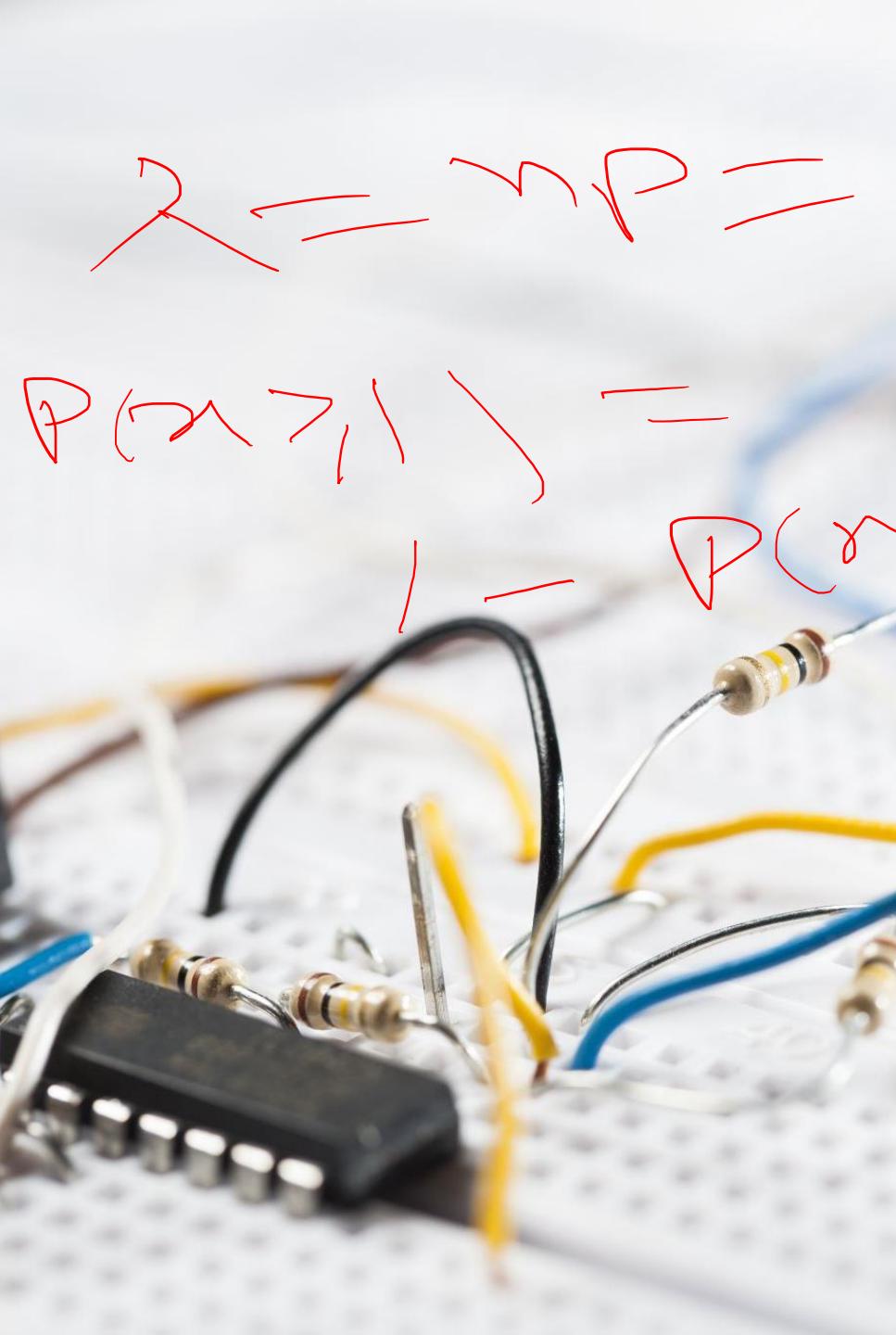
- $\text{Mean} = \lambda$

- $\text{Variance} = \lambda$

$$E(X) =$$

$$\lambda = \sum_{n=0}^{\infty} n P(X=n)$$

$$\text{Var}(X) = E(X^2) - E(X)^2$$



$$200 \times \frac{2}{100} = 4$$

## Poisson Distribution

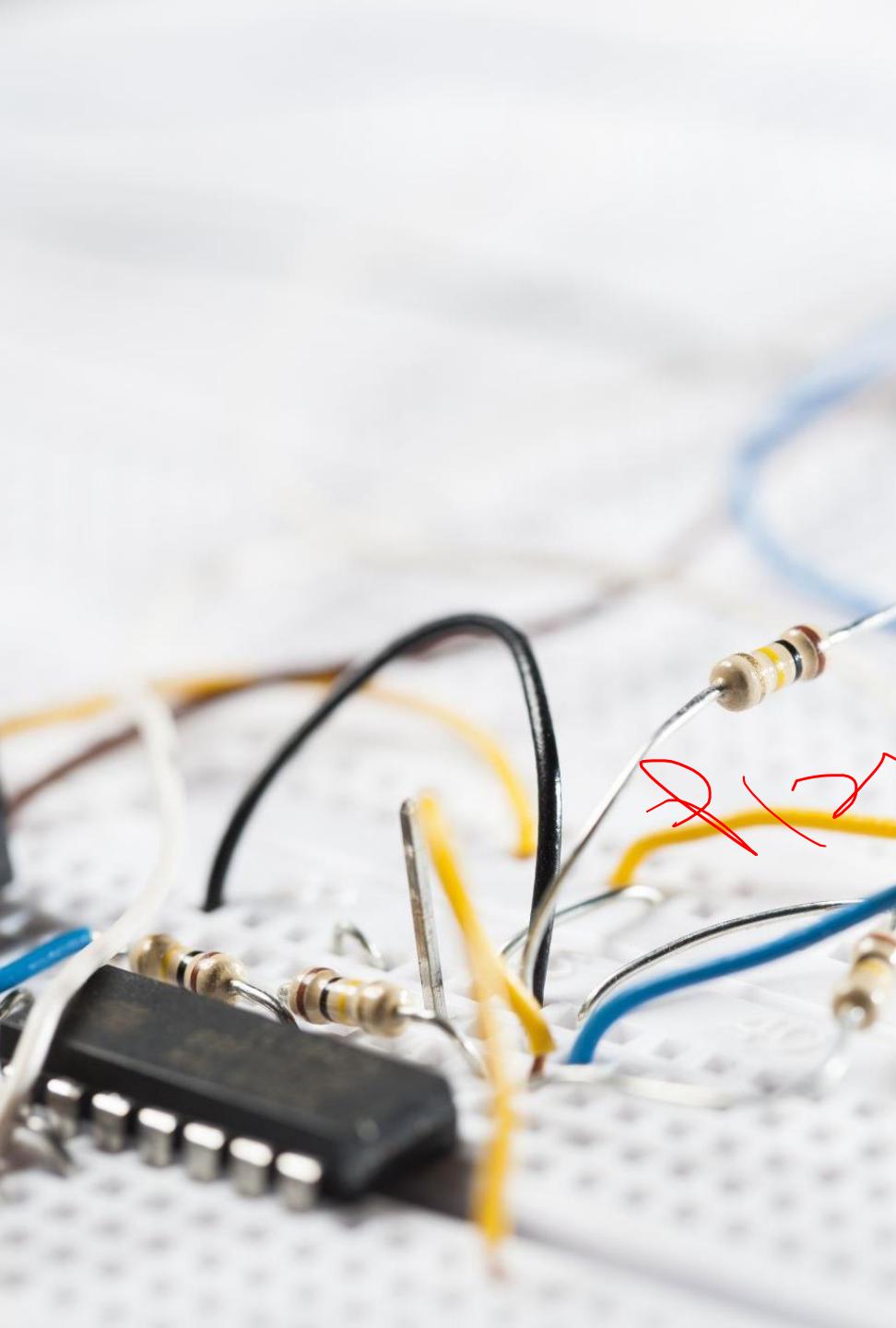
Q1. Given that 2% at the fuses manufactured by a firm are defective. Find probability that a box containing 200 fuses has

1. At least 1 defective fuses
2. 3 or more defective fuses
3. No defective fuses

$$\lambda = 4/100$$

$$n = 200$$

$$P(x > 1) ?$$



$$P(X \geq 1) = 1 - P(X=0)$$
$$= 1 - e^{-\lambda}$$

# Poisson Distribution

Solution:

$$P(X \geq 1) = 1 - P(X=0)$$
$$= 1 - e^{-\lambda}$$
$$= 1 + e^{-\lambda}$$

Diagram illustrating the Poisson distribution formula:

The top part shows a graph of the Poisson distribution function  $P(X=k)$  versus  $k$ . The curve starts at a peak and decays towards zero. A vertical arrow points from the peak to the label  $\lambda$ , and a horizontal arrow points from the peak to the label  $C$ .

The bottom part shows the calculation of  $P(X=0)$ :

$$P(X=0) = e^{-\lambda} \times C$$

A diagram below shows a horizontal line with a point labeled  $0$  on it, with a bracket underneath labeled  $C$ .

# Poisson Distribution

---

Q2. A certain factory turning out blades there is small chance of 0.02 for any blade to be defective.

The blades are supplied in packets of 10.

Find approximate number of packets containing

1. No defective blades
2. One defective blades

In Consignment of 10000 packets



# Poisson Distribution

---

Q2. A certain factory turning out blades there is small chance of 0.002 for any blade to be defective.

The blades are supplied in packets of 10. Using poisson distributions. Find approximate number of packets containing

1. No defective blades
2. One defective blades

In Consignment of 10000 packets



Solution:

# Poisson Distribution

# Poisson Distribution

---



Q3. If probability of a bad reaction from a certain infection is 0.01.



Find the chance that out of 200 individuals more than two will get bad reaction.

# Poisson Distribution

Solution:

# The Relationship Between the Binomial and Poisson Distributions

The binomial distribution tends toward the Poisson distribution as:  $n \rightarrow \infty$ ,  $p \rightarrow 0$ , and  $\lambda = np$  stays constant.

Binomial:

$$P(X = x) = \binom{n}{x} p^x (1 - p)^{n-x}$$

Poisson:

$$P(X = x) = \frac{\lambda^x e^{-\lambda}}{x!}$$

The Poisson distribution can be used to provide a reasonable approximation to the binomial distribution if  $n$  is large and  $p$  is small.

## Example:

Albinism is a rare genetic disorder that affects one in 20,000 Europeans.

People with albinism produce little or none of the pigment melanin.

$$n = 1000$$

$$p =$$

$$1/20\ 000$$

$$\lambda = 2$$

$$\lambda \approx np$$

Example:

Albinism is a rare genetic disorder that affects one in 20,000 Europeans. People with albinism produce little or none of the pigment melanin.

In a random sample of 1000 Europeans, what is the probability that exactly 2 have albinism?

# Solution

Discuss in class

$$n = \frac{2}{1000}$$

$$P = \frac{1}{2000}$$

$$0.005 = \frac{1}{2000}$$

$$\tan P = \frac{1}{2000}$$

$$\begin{aligned} n &= P \\ n &= (1 - P) \\ n &= e^{-x} \end{aligned}$$

## Solution

Binomial  $n=1000$   $p = \frac{1}{20000}$

$$P(X=2) = \binom{1000}{2} \left(\frac{1}{20000}\right)^2 \left(1 - \frac{1}{20000}\right)^{1000-2}$$
$$= \underline{0.001187965}$$

Poisson

$$\lambda = np = 1000 \cdot \frac{1}{20000} = 0.05$$

$$P(X=2) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{0.05^2 e^{-0.05}}{2!} = 0.001187$$

## *Rough* Guideline:

The Poisson approximation is reasonable if  $n > 50$  and  $np < 5$ .

# Why use this approximation?

- ▶ The factorials and exponentials in the binomial formula can become problematic to calculate.
- ▶ A problem may be binomial conceptually, but  $n$  and  $p$  may be unknown. (We may only know the mean.)

# Geometric Distribution

---

# Geometric Distribution

Geometric distribution is the probability distribution of the number  $X$  of independent Bernoulli trials performed until a success occurs, where the Bernoulli trials have a constant probability of success  $p$ .

# Geometric Distribution

## When Geometric Distribution is Used

Geometric distribution is applicable to find the probability where we perform an experiment until a success occurs.

# Geometric Distribution

## Examples:

- 1) Tossing a coin repeatedly until the first head appears.
- 2) Shot the target until it hits.
- 3) Give the test until he will pass it.
- 4) Throwing a die repeatedly until first time a six appears.

# Geometric Distribution

$$P(X=10) = q^9 \cdot p^1$$

**Example:** Suppose that a trainee solder shoots a target in an independent fashion. If the probability that the target is hit in any shot is 0.7, what is the

$$= 0.3^9 \cdot 0.7$$

- i) Probability that the target would be hit on 10<sup>th</sup> attempt?

**Solution:** Required probability

$$p = 0.7$$

$$q = 0.3$$





# Geometric Distribution

**Example:** Suppose that a trainee solder shoots a target in an independent fashion. If the probability that the target is hit in any shot is 0.7, what is the

- ii) Probability that it takes him less than 4 shots?

**Solution:** Required probability

$$\begin{aligned} P(X < 4) &= P(1) + P(2) + P(3) \\ &= p + qp + q^2p \end{aligned}$$

Bion.

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

$$\forall e = P(X) = [{}^{n-1}C_{x-1} p^{x-1} q^{(n-1)-(x-1)}] p$$

$$l = hp$$

Geom =  $P(X) = q^{X-1} \cdot p$

$$P(X) = \frac{\gamma^X \cdot e^{-\gamma}}{x!}$$

PoiM =

$$= \sqrt{p} [1 + q\sqrt{2} + q^4 + \dots]$$

Geometric Distribution

$$= \sqrt{p} \frac{1}{1 - q\sqrt{2}} = 0.2368$$

**Example:** Suppose that a trainee solder ~~shoots~~ a target in an independent fashion. If the probability that the target is hit in any shot is 0.7, what is the

iii) Probability that it takes him an even number of shots?

**Solution:** Required probability

$$\begin{aligned} P_{\text{even}} &= P(2) + P(4) + P(6) + \dots \\ &= q\sqrt{p} + q^3 p + q^5 p + \dots \end{aligned}$$





# Geometric Distribution

Example 2: Suppose that the probability for an applicant for a driver's licence to pass the road test on any given attempt is  $\frac{2}{3}$ . What is the probability that the applicant will pass the road test on the third attempt?

$$p = \frac{2}{3}$$
$$q = \frac{1}{3}$$

Pass  
fail

$$P(X=3) = q^{x-1} \cdot p$$
$$= \frac{1}{3}^2 \cdot \frac{2}{3} =$$



# Geometric Distribution

**Example:** Suppose that a trainee solder shoots a target according to a geometric distribution. If the probability that a target is shot in any shot is 0.8, find the probability that it takes an Odd number of shots.

**Example:** Let  $X$  be the number of births in a hospital until the first girl is born. Determine the probability and the distribution function of  $X$ .

Assume that probability that the baby born is a girl is  $\frac{1}{2}$ .

Geometrische Dif?: Ex)  $\frac{x^2}{x+1}$  Bionomial Dif

mean  $\Rightarrow$  // p

$$\text{Varianz} \Rightarrow \frac{\sigma^2}{\mu^2}$$

S. I →

$$P(x) = 2^{x-1} \cdot \bar{c}$$

$$mam = \uparrow \nwarrow$$

$$\text{Var} = r + v$$

S.D =

$C \times \beta^+$

Watt-  $\frac{γ}{p}$

$$v_{\text{air}} = \frac{r}{t}$$

A hand-drawn diagram illustrating the relationship between  $S \cdot D$  and  $r_m$ . At the top left, the expression  $S \cdot D$  is written above a horizontal line. To the right of the line, there is an equals sign followed by a vertical line segment labeled  $r_m$ . Below the horizontal line, there is a curved arrow pointing upwards and to the right, originating from a point on the line. The label  $x^-$  is placed near the start of this arrow. To the right of the arrow, the letter  $e$  is written vertically. Below the arrow, the label  $x^+$  is placed near its end.

# Geometric Distribution

- Geometric distribution is a type of discrete probability distribution that represents the probability of the number of successive failures before a success is obtained in a Bernoulli trial.
- In other words, in a geometric distribution, a Bernoulli trial is repeated until a success is obtained and then stopped.

## Example:

- Cost-Benefit Analysis
- Sports Applications

The probability that a batter or hitter is able to make a successful hit before three strikes can be estimated efficiently with the help of a geometric probability distribution function.

# Geometric Distribution

## Assumptions:

- The trials being conducted are independent.
- There can only be two outcomes of each trial - success or failure.
- The success probability, denoted by  $p$ , is the same for each trial.

## Example:

Suppose a dice is repeatedly rolled until "3" is obtained. Then the probability of getting "3" is  $p = 1 / 6$  and the random variable,  $X$ , can take on a value of 1, 2, 3, ...., until the first success is obtained. This is an example of a geometric distribution with  $p = 1 / 6$ .

# Geometric Distribution

Geometric Distribution PMF

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

Geometric Distribution CDF

$$P(X \leq x) = 1 - (1 - p)^x$$

# Mean and Variance of Geometric Distribution

- $E[X] = 1 / p$
- $Var[X] = (1 - p) / p^2$

# Example

- If your probability of success is 0.2, what is the probability you meet an independent voter on your third try?

Sol:

$$p = 0.2$$

$$X = 3$$

$$f(x) = (1 - p)^{x-1} * p$$

$$P(X = 3) = (1 - 0.2)^{3-1}(0.2)$$

$$P(X = 3) = (0.8)^2 * 0.2 = 0.128$$

# Example

- Suppose Rohit owns a pen manufacturing company and determines that 3 out of every 75 pens are defective.

What is the probability that Rohit will find the first faulty pen on the 6th one that he tested?

**Sol:**

$$p = 3/75 =$$

$$X = 6$$

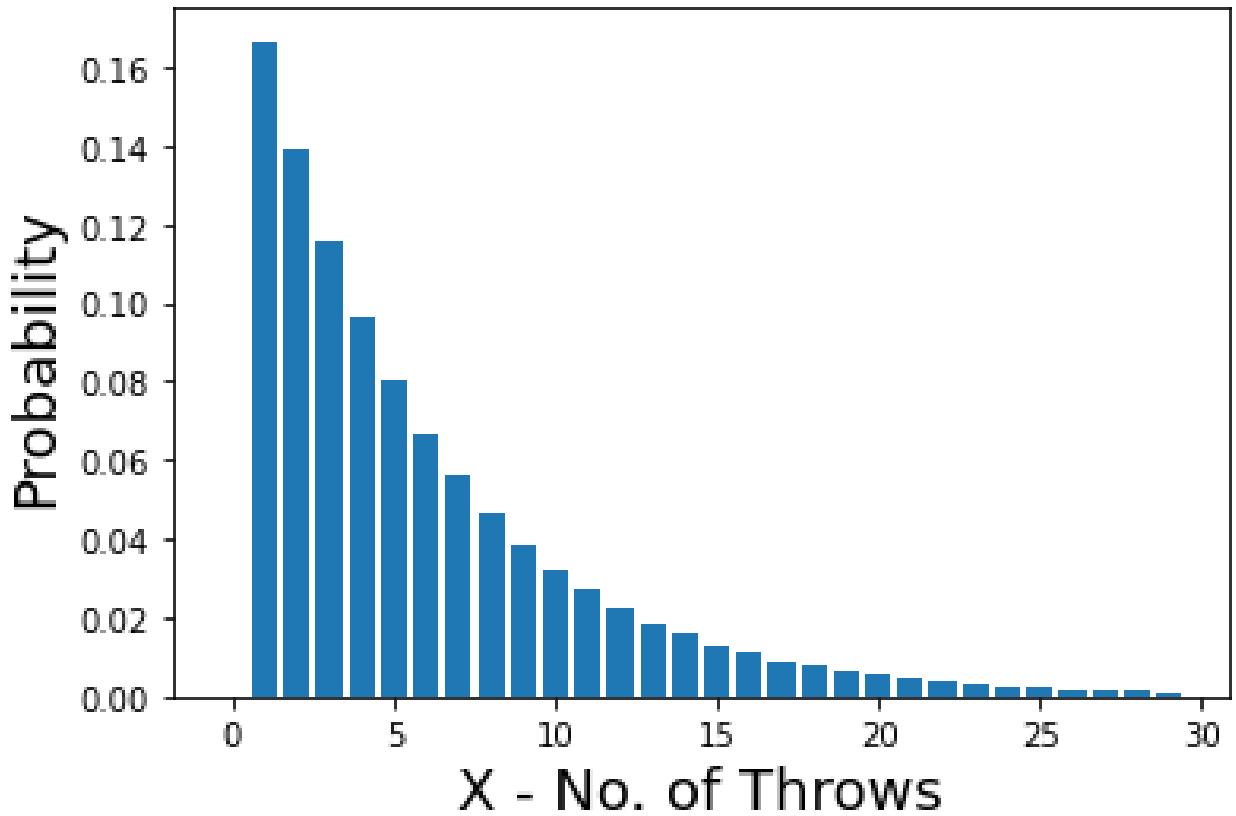
$$f(x) = (1 - p)^{x-1} * p$$

$$P(X = 6) = (1 - 3/75)^{6-1}(3/75)$$

$$P(X = 6) = (0.96)^5 * (0.04) =$$

# Geometric Distribution

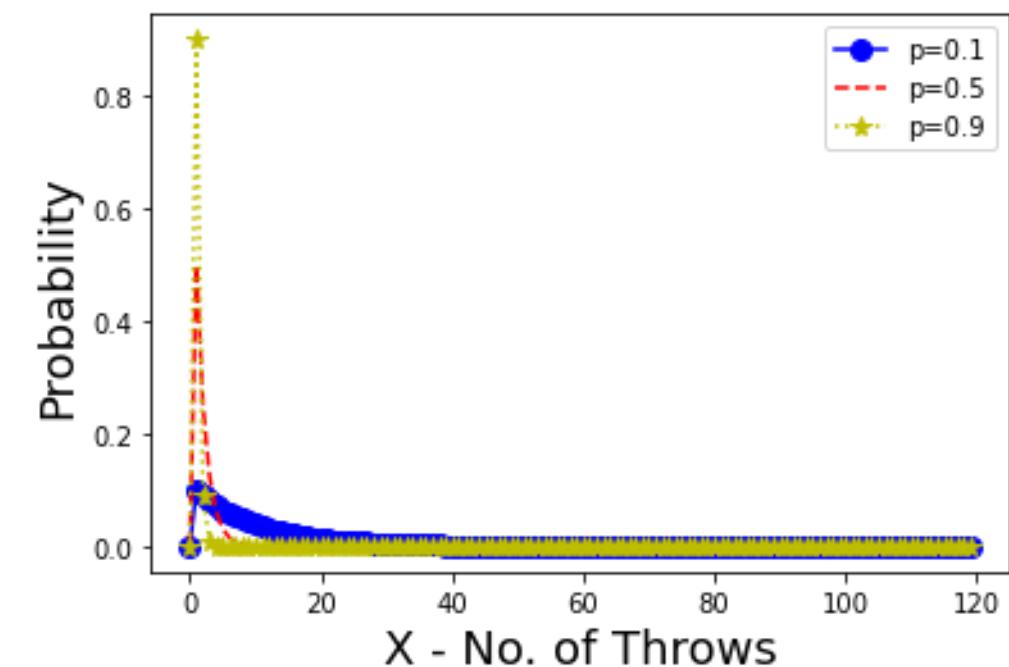
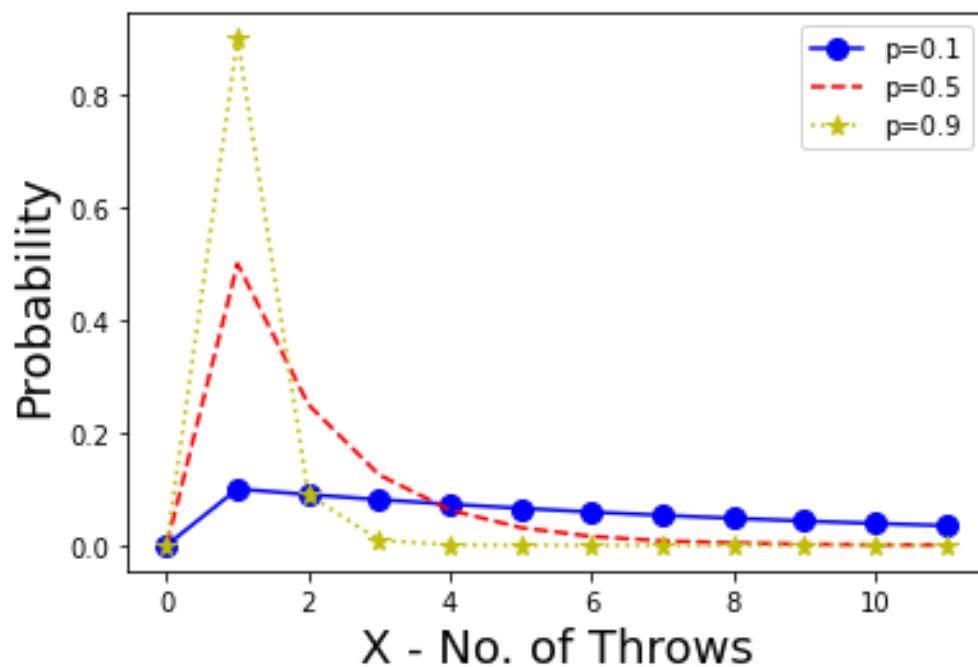
- The geometric distribution graph below displays the probability of rolling the first six in precisely 1, 2, 3, etc. rolls, up to 30.



Source: My Python Example

# Geometric Distribution

Source: My Python Example



# Geometric Distribution vs Geometric Series

Geometric Distribution PMF

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

Now, put  $x = 1, 2, 3, \dots$

We get,

$$p, (1 - p)p, (1 - p)^2 p, (1 - p)^3 p \dots \dots \dots$$

Which is a geometric series.....

Sum = 1 (see python example)



# Continuous random variables

# Outline

---

Discrete vs continuous random variables

---

Probability mass function vs Probability density function

---

Properties of the pdf

---

Cumulative distribution function

---

Properties of the cdf

---

Expectation, variance and properties

---

The normal distribution

Till now, we discussed

Discrete random variables: can take a finite, or at most countably infinite, number of values,

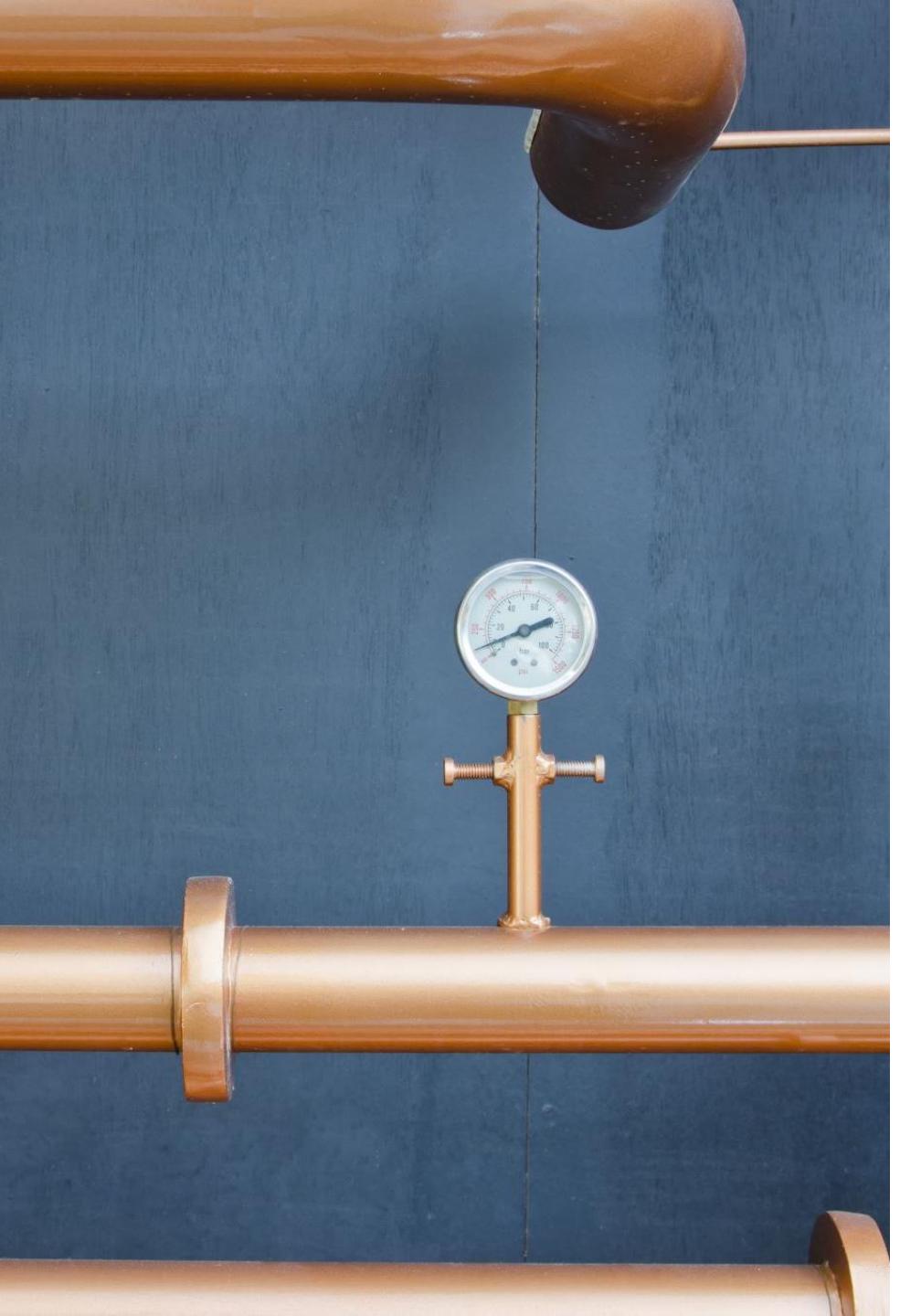
For example:

- Binomial random variable
- Bernoulli random variable
- Geometric random variable

# Continuous random variable

- A continuous random variable is a random variable that can take on an uncountably infinite range of values.
- Due to the above definition, the probability that a continuous random variable will take on an exact value is 0.
- For any specific value  $X = x$ ,  $P(X = x) = 0$

*fx*

A photograph of a section of shiny copper pipe. A small, round pressure gauge is mounted vertically onto the pipe. The gauge has a white face with black markings and numbers, including 0, 20, 40, 60, 80, 100, and 120. The needle is positioned between 60 and 80. The pipe is set against a dark, textured background.

# Continuous random variable

## Examples:

- The volume of water passing through a pipe over a given time period.
- The height of a randomly selected individual.

$$P(x) = \sum p_i$$

$$P(x) = \int_{-\infty}^{\infty} f(x) dx$$

$$\int_0^1 4x^3 dx$$

Example:

Suppose the **probability density function** of a continuous random variable, X, is given by  $4x^3$ , where  $x \in [0, 1]$ .

The probability that X takes on a value between  $1/2$  and  $1$  needs to be determined.

$$f(x) = \begin{cases} 4x^3 & 0 \leq x \\ 0 & \text{elsewhere} \end{cases}$$

# Solution

**Solution:**

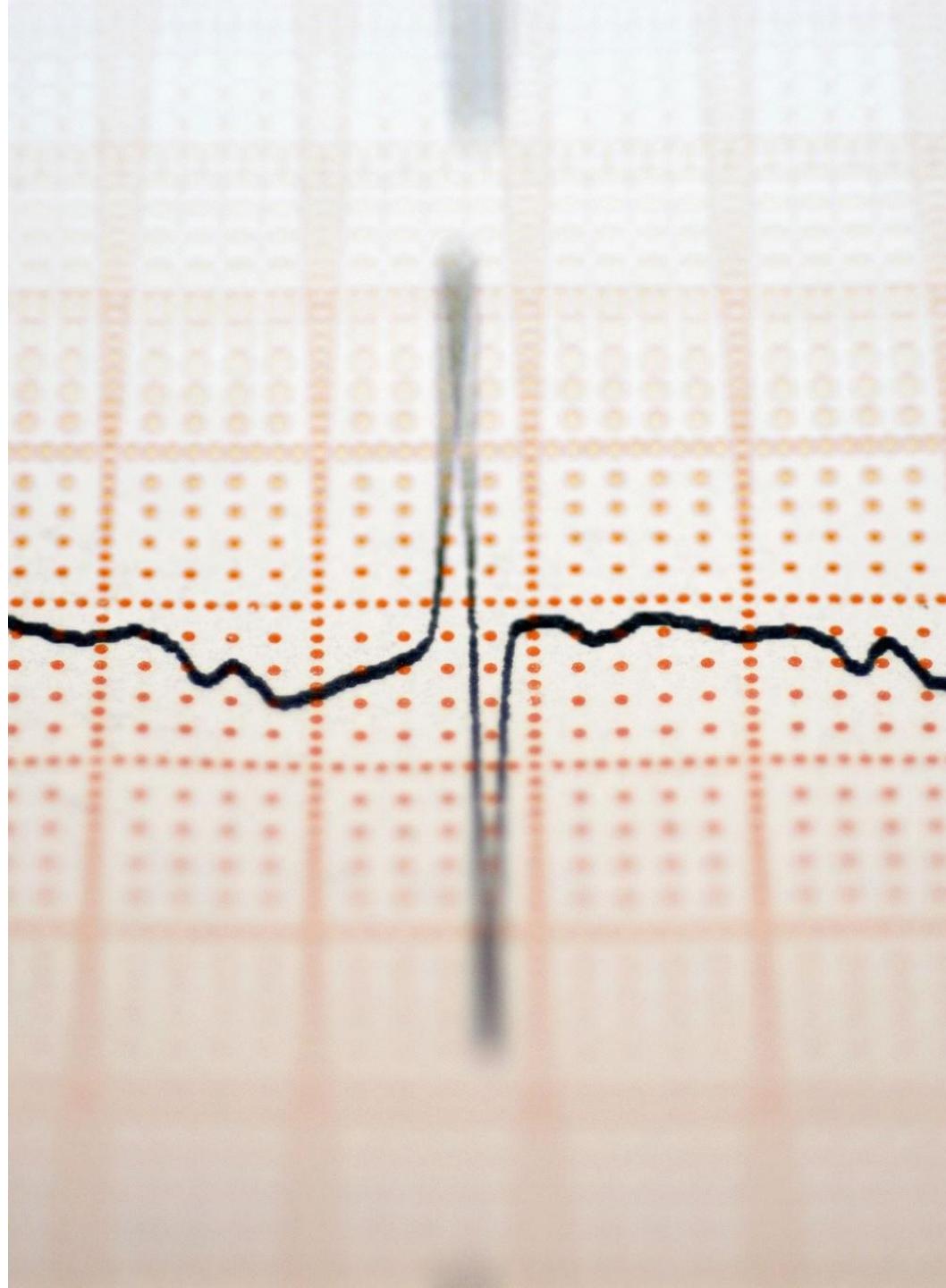
This can be done by integrating  $4x^3$  between  $1/2$  and  $1$ . Thus, the required probability is

$$P = \int_{1/2}^1 4x^3$$

$$P = 15/16.$$

# Probability density function (pdf)

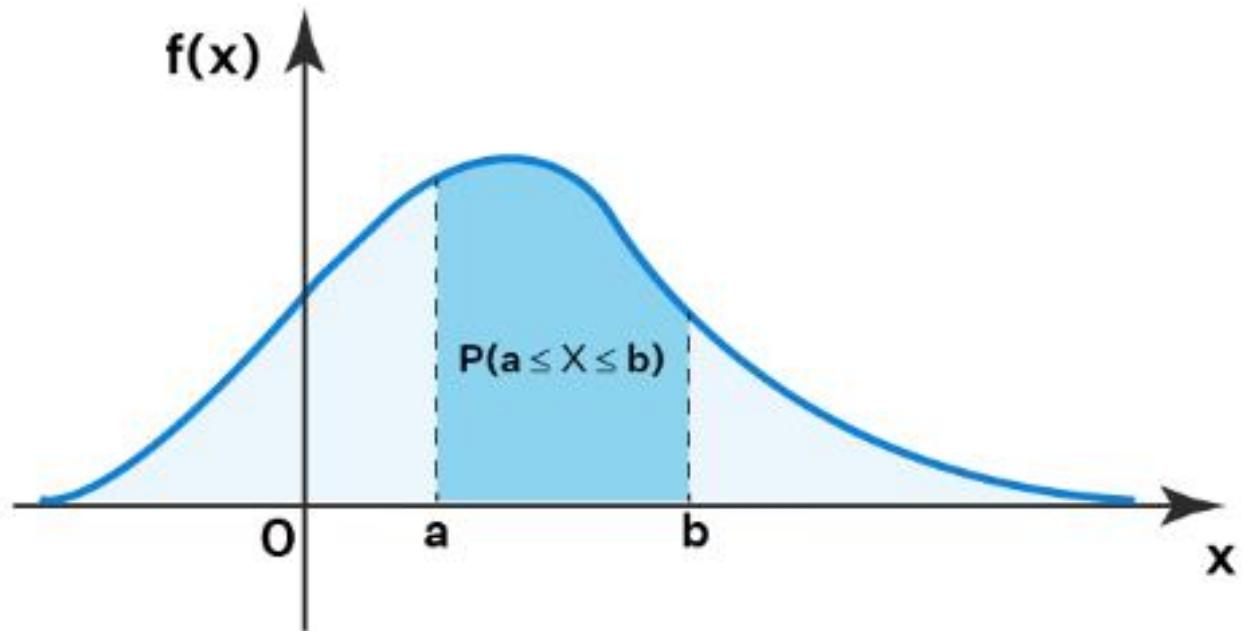
- The probability density function (pdf) and the cumulative distribution function (CDF) are used to describe the probabilities associated with a continuous random variable.
- For a continuous random variable, we cannot construct a PMF (discussed earlier for discrete) – each specific value has zero probability.
- Instead, we use a continuous, non-negative function  $f_X(x)$  called the probability density function, or PDF, of  $X$



## Probability density function (pdf)

The probability of  $X$  lying between two values  $a$  and  $b$  is simply the area under the PDF, i.e

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$



# Example

The pdf of a continuous random variable, X, is given as follows:

$$f(x) = \begin{cases} x & 0 \leq x \leq 1 \\ x + 3 & 1 < x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Find the probability that the value of the continuous random variable will lie between 0 and 0.5, i.e., Find  $P(0 \leq X \leq 0.5)$ .

Ans: 0.125

$$P(0 \leq X \leq 0.5) = \int_0^{0.5} x^2 dx$$

# Solution

-

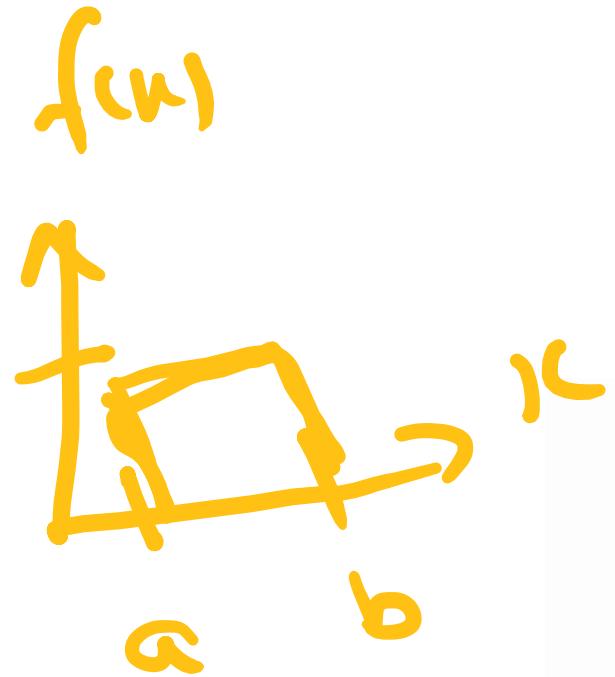
$$P(X) = \sum P(x) = 1$$

## Properties of the pdf

- For any single value  $a$ ,  $P(X = a) = \int_a^a f_X(x)dx = 0$
- $f(x) \geq 0$ . This implies that the probability density function of a continuous random variable cannot be negative.
- $\int_{-\infty}^{+\infty} f_X(x)dx = 1$ , this means that the total area under the graph of the pdf must be equal to 1.
- Note that  $f(x)$  can be greater than 1 – even infinite! – for certain values of  $x$ , provided the integral over all  $x$  is 1.



# Uniform Distribution



Constant Probability

within Domain

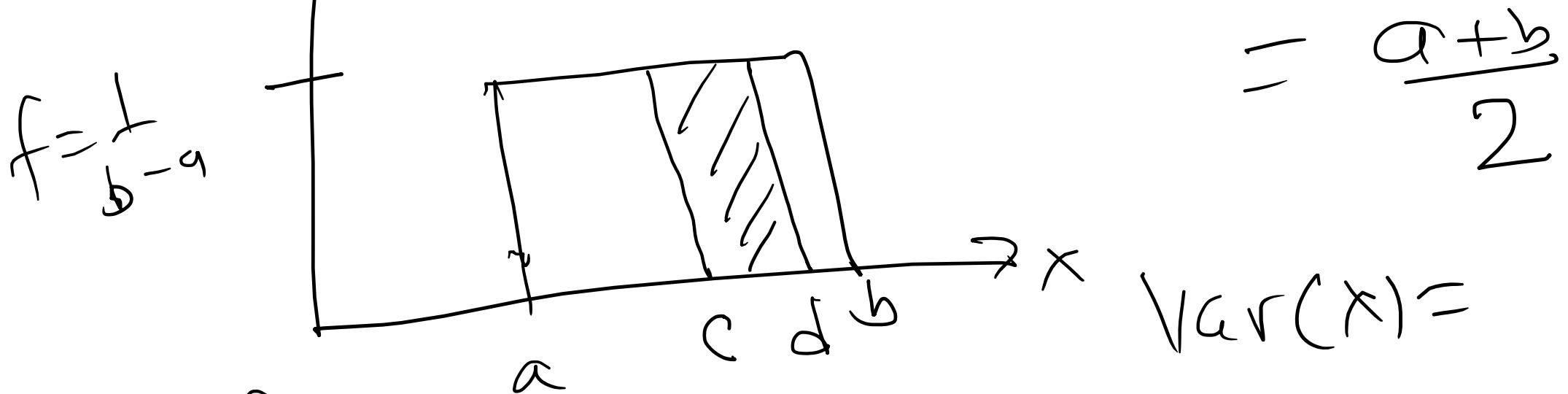
$$f(x) = \begin{cases} K = \frac{1}{b-a} & a < x < b \\ 0 & \text{Otherwise} \end{cases}$$

P.D.F. =  $\int_a^b f(x) dx = \int_a^b K dx = K \cdot (b-a)$

$f(x) = \frac{1}{b-a}$

$\int_a^b f(x) dx = \int_a^b K dx \Rightarrow K \cdot (b-a) = 1 \Rightarrow K = \frac{1}{b-a}$

$$f_{Cu} \uparrow P(c \leq x \leq d) = \frac{d-c}{b-a} \text{ aus } E(x)$$



$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_a^b x \cdot \frac{1}{b-a} dx = \frac{1}{b-a} \int_a^b x dx$$

$$\Rightarrow \frac{1}{b-a} \int_a^b x \, dx = \frac{1}{b-a} \left. \frac{x^2}{2} \right|_a^b$$

$$\Rightarrow \frac{1}{b-a} \frac{\frac{1}{2} (b-a)}{2} = \frac{1}{b-a} \frac{(a+b)(b-a)}{2}$$

$$\Rightarrow \frac{1}{2} (a+b)$$

$$\text{Var}(X) = E(X^2) - (E(X))^2$$

$$\text{Var}(x) = E(x^2) - (E(x))^2$$

$$E(x^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx \Rightarrow \frac{1}{b-a} \int_a^b x^2 dx$$

$$= \frac{x^3}{3} \Big|_a^b \Rightarrow \frac{1}{b-a} \frac{b^3 - a^3}{3}$$

~~$(b-a)$~~   $\cancel{(b-a)}(a^2 + ab + b^2)$

$$\Rightarrow \frac{a^3 + b^3 + ab}{3}$$

$$\text{Var}(X) = \frac{a^2 + b^2 + ab}{3} - \frac{(a+b)^2}{4}$$

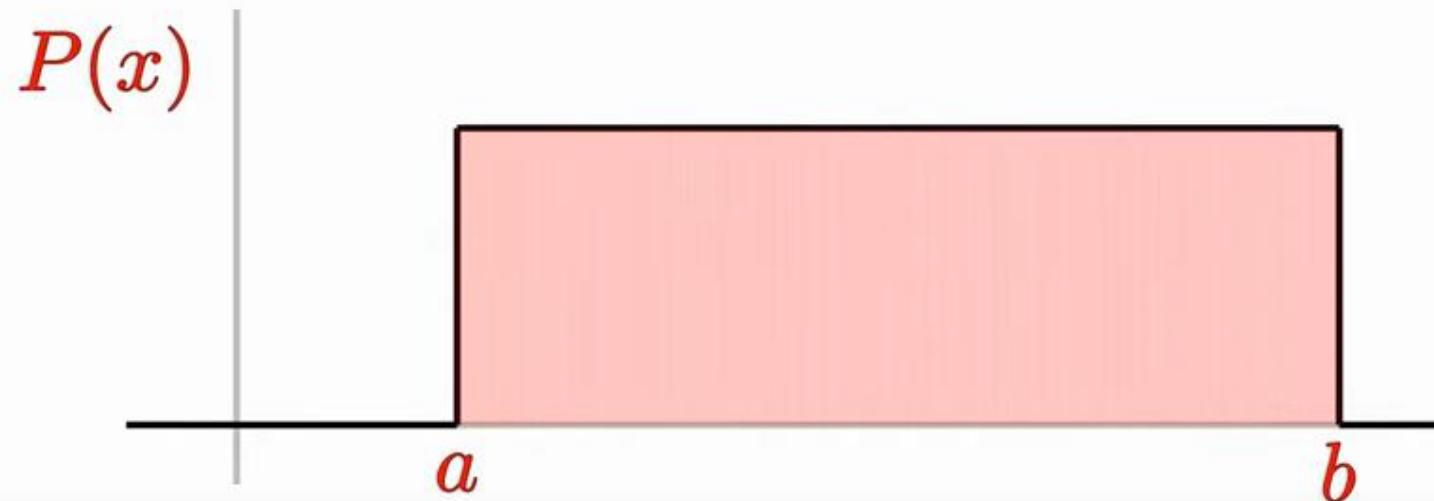
$$= \frac{4(a^2 + b^2 + ab) - 3(a^2 + b^2 + 2ab)}{12}$$

$$= \frac{a^2 + b^2 - 2ab}{12} = \frac{(a-b)^2}{12}$$

$$E(X) = \frac{a+b}{2};$$

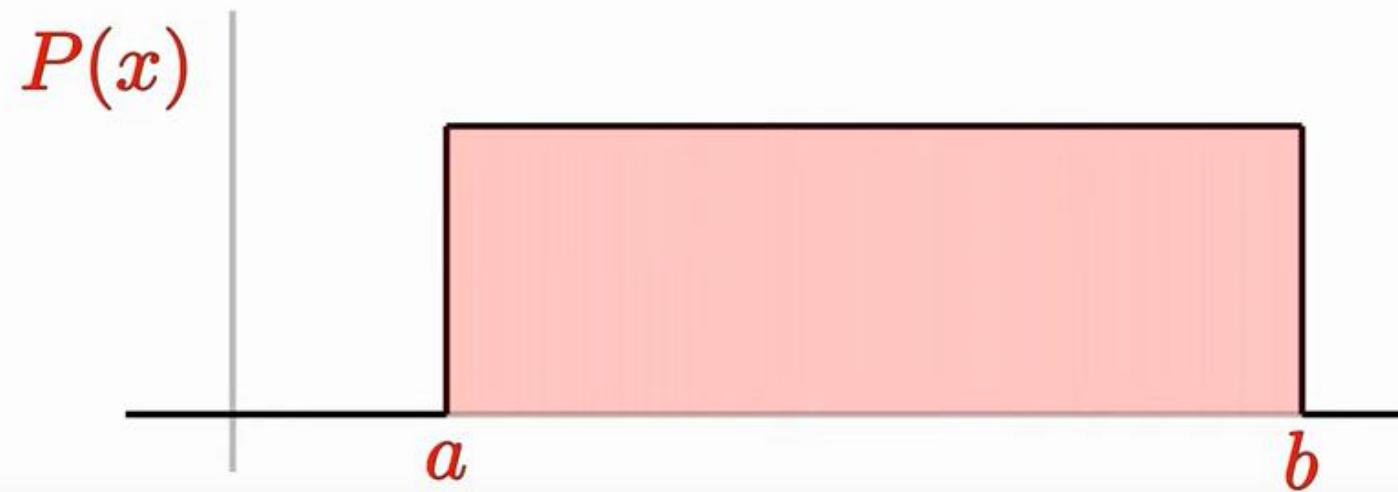
$$S \cdot D = \sqrt{\frac{(a-b)^2}{12}}$$

# Uniform Distribution



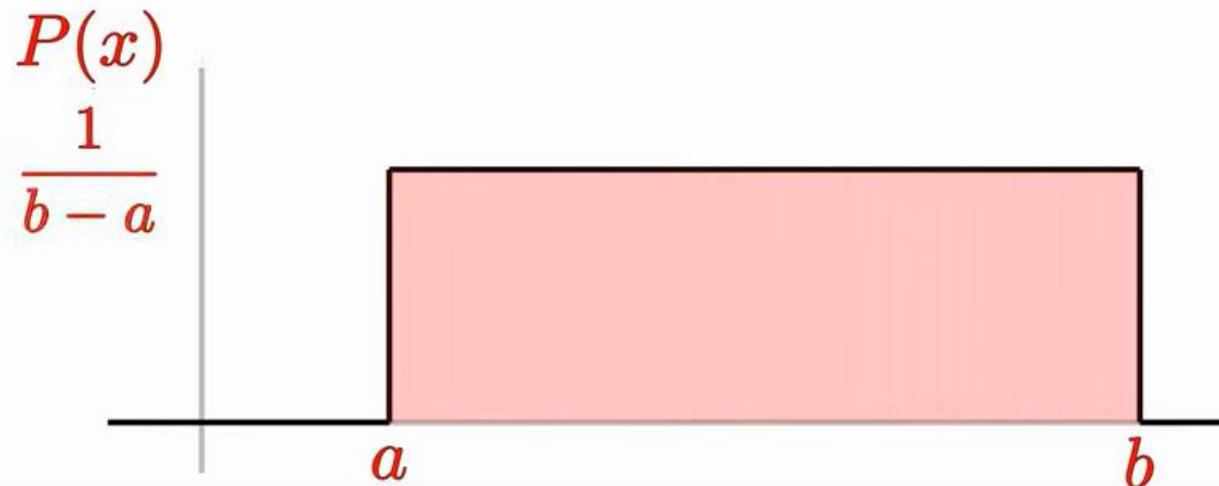
$$P(a \leq X \leq b) = 1$$

# Uniform Distribution



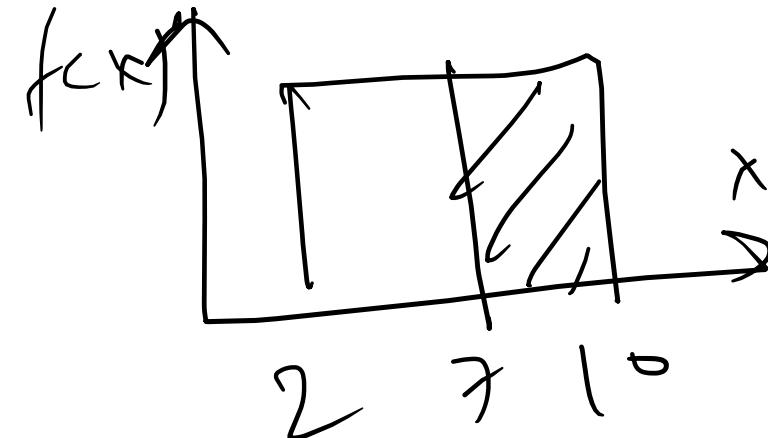
$$P(a \leq X \leq b) = 1 = \text{height} \cdot \text{width}$$

# Uniform Distribution



$$P(a \leq X \leq b) = \frac{1}{(b-a)} = \text{height} \cdot \frac{(b-a)}{\cancel{(b-a)}} \Rightarrow \text{height} = \frac{1}{b-a}$$

# Uniform Distribution

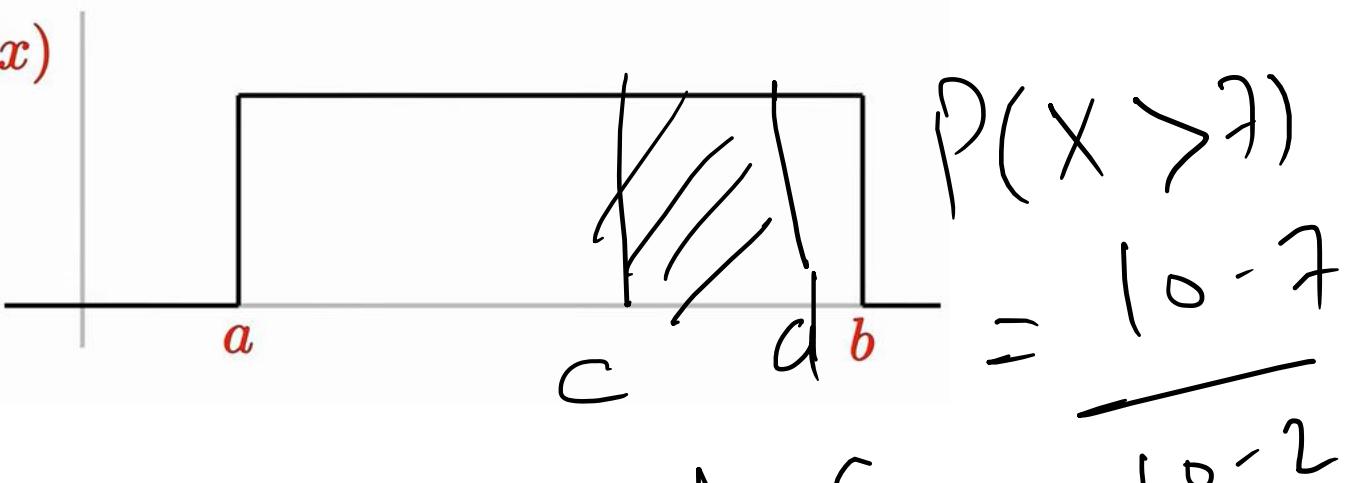


$$\mu = \frac{a+b}{2}$$

$$\sigma = \sqrt{\frac{(b-a)^2}{12}}$$

$$P(c \leq X \leq d) = \frac{d-c}{b-a}$$

$$P(x)$$



$$\frac{d-c}{b-a} = \frac{10-7}{10-2}$$



$$= 0.375$$

$$= \frac{3}{8}$$

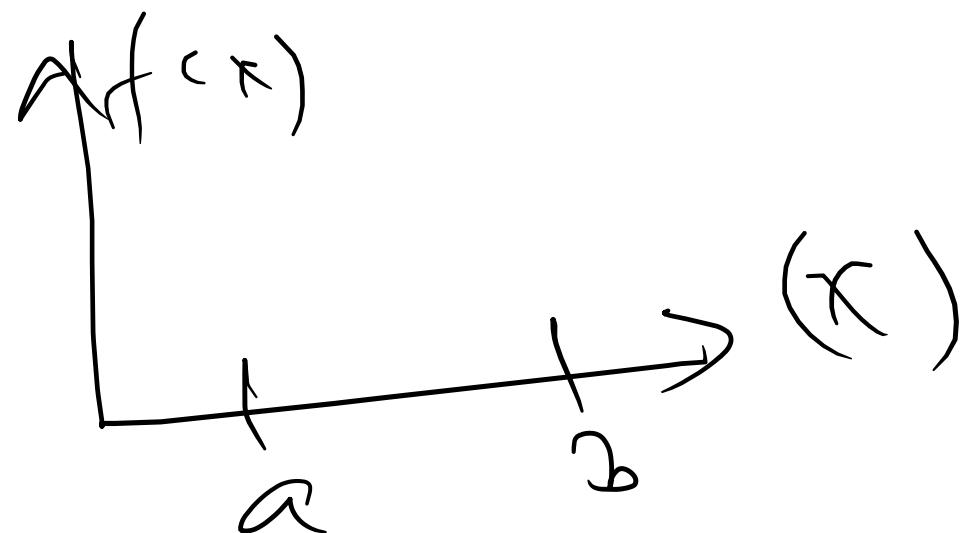
Q:- If  $X$  is uniformly distribution  
with mean  $\Rightarrow 1$  and variance  $= 4$

Find  $P(X < 0) = ?$

$$P(X) = \int_a^{\infty} f(x) dx$$

$$E(X) = 1$$

$$\text{Var}(X) = 4$$



meerv = 1:

$$\frac{a+b}{2} = 1$$

$$a+b = 2$$

$$\begin{aligned} a-b &= \pm 4 \\ a+b &= \end{aligned} \quad \left. \begin{array}{l} +4 \\ -4 \end{array} \right\} \rightarrow \textcircled{+}$$



$$\begin{aligned} a-b &= -4 \\ a+b &= 2 \end{aligned} \quad \left. \begin{array}{l} -4 \\ +2 \end{array} \right\} \rightarrow \textcircled{-}$$

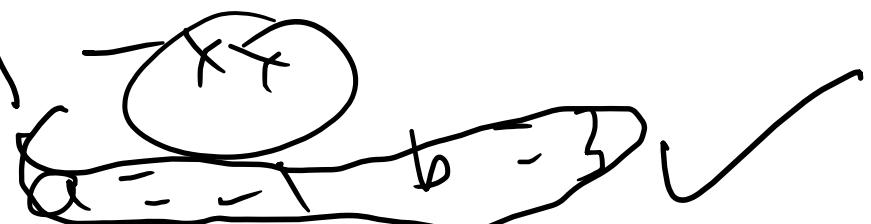
VG v(x) = 4/3

$$\frac{(a-b)^2}{12} = 4/3$$

$$(a-b)^2 = \frac{4 \times 12}{3}$$

$$(a-b)^2 = 16$$

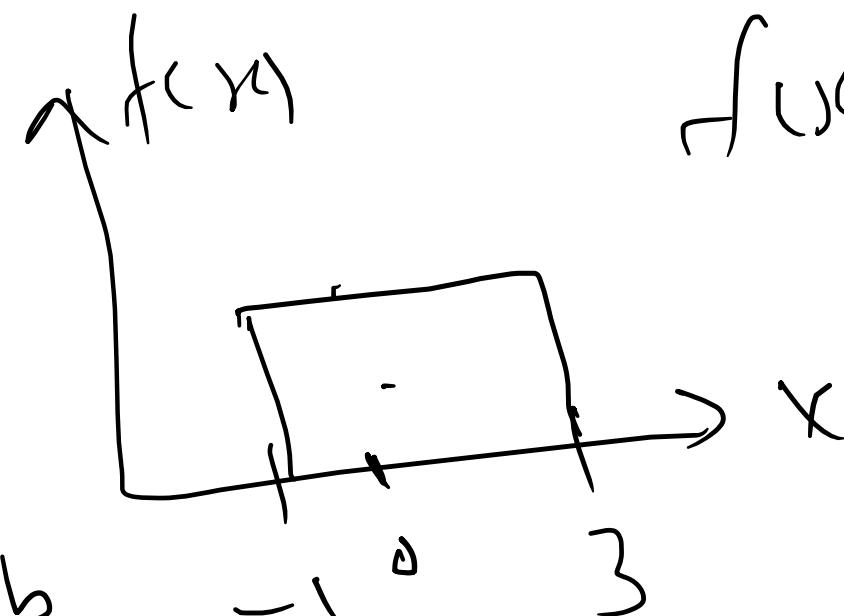
$$a-b = \pm 4$$



$$a = -1$$

$$b = 3$$

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$



$$\int f(x) dx = 1/4$$

$$f(x) = \begin{cases} \frac{1}{3 - (-1)} & -1 < x < 3 \\ 0 & \text{otherwise} \end{cases}$$

$$P(X < 0) = \int_{-1}^0 f(x) dx \quad \left| \begin{array}{l} \frac{1}{4}x \\ -1 \end{array} \right|^0$$

$$= \int_{-1}^0 f(x) dx \Rightarrow \frac{1}{4}$$

$$= \int_{-1}^0 1/4 dx$$



$$F(x) = \frac{a+b}{2} = \frac{12}{2} = 6$$

## Uniform Distribution

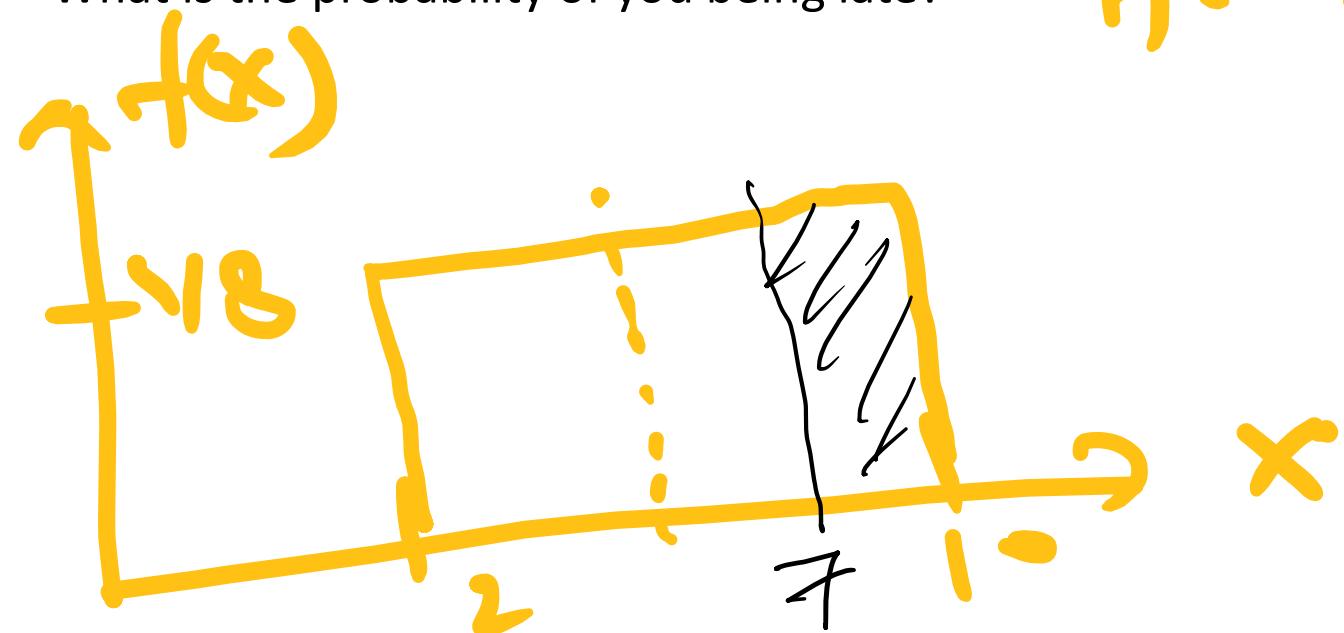
S.D =

Ex. Bus is uniformly late between 2 and 10 minutes. How long can you expect to wait? With S.D ?

If its >7 mins late, you will be late for work.

What is the probability of you being late?

$$f(x) = \frac{1}{b-a}$$



$$\sqrt{a(*)} = \sqrt{(a - b)^2}$$

$$= \sqrt{(2 - 10)^2} =$$

$$\begin{aligned} a &= 2 \\ b &= 10 \\ \frac{64}{12} &= 2.31 \end{aligned}$$