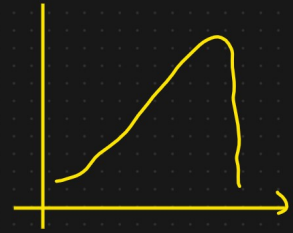
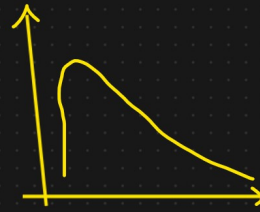
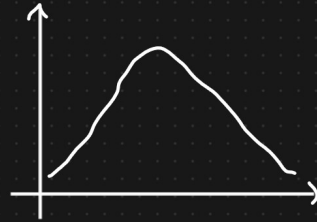
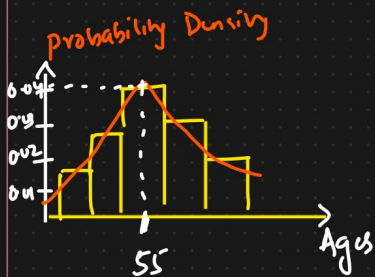


# Probability Distribution Functions

Probability distribution functions describe how the probabilities are distributed over the values of a random variable.

$Agus = \{ - - - - - \} \Rightarrow$  Continuous random variable



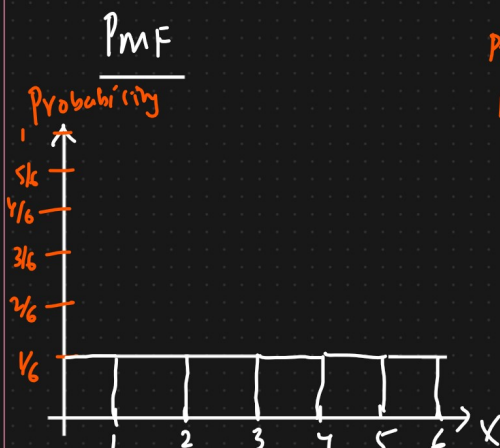
2 Main of probability distribution functions

- ① Probability Mass functions (PMF) : Used for discrete random variables.
- ② Probability Density functions (PDF) : Used for continuous random variables.
- ③ Cumulative Distribution function (cdf).

① Probability Mass Function [Discrete Random Variable]

Eg: Rolling a dice  $X = \{1, 2, 3, 4, 5, 6\} \Rightarrow$  Fair Dice

$$Pr(1) = Pr(2) = Pr(3) = Pr(4) = Pr(5) = Pr(6) = \frac{1}{6}$$



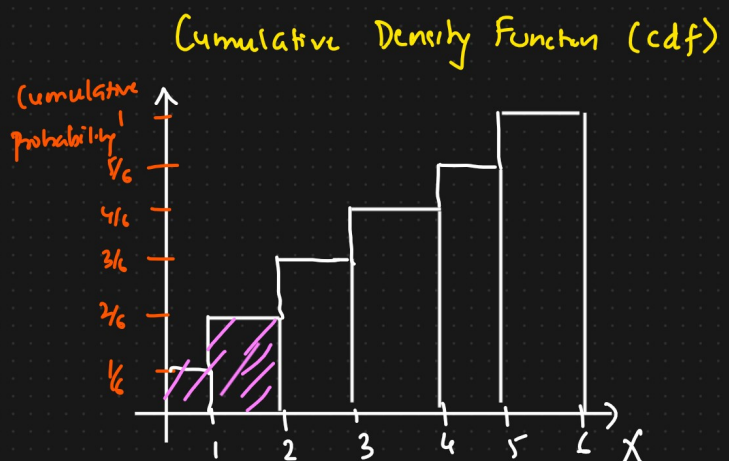
$$Pr(1) = \frac{1}{6}$$

$$Pr(2) = \frac{1}{6}$$

$$Pr(3) = \frac{1}{6}$$

$$\vdots$$

$\Rightarrow$



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

$$\begin{aligned} Pr(X \leq 6) &= Pr(X=1) + Pr(X=2) \\ &\quad + \dots + Pr(X=6) \\ &= 1 \end{aligned}$$

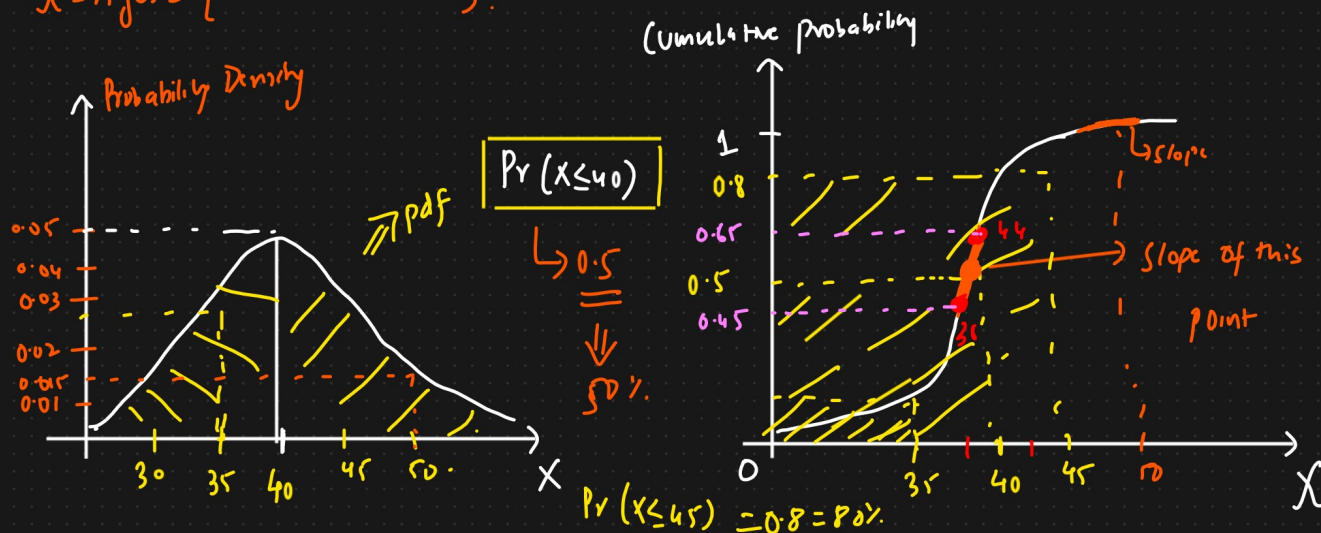
## ② Probability Density Function (pdf)

① Distribution of Continuous Random Variable

① Area under the curve ✓

② Probability density ✓

$$X = \text{Age} = \{ \dots \}$$



$(x_1, y_1)$

$(x_2, y_2)$

$$\text{Slope} = \left[ \frac{y_2 - y_1}{x_2 - x_1} \right] \leftarrow \text{Gradient} \Rightarrow \text{probability density}$$

$\text{Slope} = \left\{ \frac{0.65 - 0.45}{44 - 38} \right\} \Rightarrow \text{probability density}$

Probability Density = Gradient of Cumulative Density function

### PDF Properties

① Non Negativity  $f(x) \geq 0$  for all  $x$

② The total area under the PDF curve is equal to 1

$$\int_{-\infty}^{\infty} f(x) dx = 1$$



With respect to different distribution

$f(x)$  function is going to change

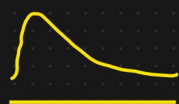




# Different distribution types

## Types of Probability Distribution [pdf, pmf, cdf]

Age, Weight, Salary



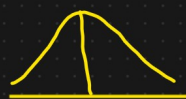
DATASET  $\Rightarrow$  Distribution

SETS OF  $\Downarrow \Uparrow$

① Bernoulli Distribution  $\rightarrow$  Outcomes are binary (pmf)  $\Rightarrow$  Discrete Random Variable

② Binomial Distribution  $\rightarrow$  (pmf)

③ Normal/Gaussian Distribution  $\rightarrow$  (pdf)  $\Rightarrow$



$\Rightarrow$  Assumptions.

④ Poisson Distribution (pmf)

⑤ Log Normal Distribution (pdf)

⑥ Uniform Distribution (pmf)

Dataset  $\rightarrow$  House price prediction Dataset-

[EDA, FE]  $\Rightarrow$  DATA ANALYST  
DATA SCIENTIST

Size of the house	No. of Rooms	Location	Floor	Sea Side	Price.
$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	$\downarrow$	
Continuous random Variable	{ Discrete }		{ Discrete }	{ <u>0 &amp; 1</u> }	Continuous
$\downarrow$			pmf	pmf	pdf
pdf.					